DOCUMENT RESUME

ED 035 121	40	EC 004 361

AUTHOR	Warner, Dennis A.; de Jung, John E.
TITLE	Goal Setting Behavior as an Independent Variable Related to the Performance of Educable Mentally Retarded Male Adolescents on Educational Tasks of Varying Difficulty. Final Report.
INSTITUTION	Washington State Univ., Pullman.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
BUREAU NO	BR-7-I-115
PUB DATE	Mar 69
GRANT	OEG-9-8-070115-0080(032)
NOTE	76p.
EDRS PRICE	EDRS Price MF-$0.50 HC-$3.90
DESCRIPTORS	Adolescents, Aspiration, Educable Mentally Handicapped, *Exceptional Child Research, Feedback, *Goal Orientation, Institutionalized (Persons), Males, *Mentally Handicapped, Motivation, Spelling, Task Performance, Testing

ABSTRACT

To investigate the effects of goal setting upon the spelling performance of educable mentally retarded male adolescents, comparisons were made between institutionalized and non-institutionalized subjects on the effects of goal setting and of task difficulty upon performance. A pilot study developed appropriate educational tasks for use in later testing with standardization of instructions and experimental procedures. Forty institutionalized and 40 special education educable mentally retarded male adolescents, randomly selected from among the 93 pilot study subjects, were assigned to one of the following four treatment groups: easy task-knowledge of past performance plus statement of goal, hard task-knowledge of past performance plus statement of goal, easy task-knowledge of past performance only, and hard task-knowledge of past performance only. The 80 subjects were tested alone on the spelling tasks by a single experimenter. Results indicated that higher scores were associated with goal setting than with feedback information only, and that higher scores were associated with the easy task rather than with the hard task. The expectation that goal setting would be superior to non-goal setting in terms of performance on a hard task but not on an easy task was upheld for the public school sub-sample but not for the institutionalized sub-sample. (Author/LE)

ED035121

BR 7-I-115
OE-BR

PA 40

Final Report

Project No. 7-I-115
Grant No. OEG 9-8-070115-0080(032)

GOAL SETTING BEHAVIOR AS AN INDEPENDENT VARIABLE RELATED TO THE PERFORMANCE OF EDUCABLE MENTALLY RETARDED MALE ADOLESCENTS ON EDUCATIONAL TASKS OF VARYING DIFFICULTY

Dennis A. Warner
Washington State University
and
John E. de Jung
University of Oregon

March, 1969

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE
Office of Education
Bureau of Research

EC 004 361E

ED035121

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Final Report

Project No. 7-I-115
Grant No. OEG 9-8-070115-0080(032)

GOAL SETTING BEHAVIOR AS AN INDEPENDENT VARIABLE RELATED TO THE PERFORMANCE OF EDUCABLE MENTALLY RETARDED MALE ADOLESCENTS ON EDUCATIONAL TASKS OF VARYING DIFFICULTY

Dennis A. Warner
Washington State University
and
John E. de Jung
University of Oregon

1969

The work presented or reported herein was performed pursuant to a grant from the United States Office of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred.

ACKNOWLEDGMENTS

The writers wish to express appreciation to the following research consultants for guidance received in developing, conducting, and/or reporting the present study: Drs. Knute Espeseth, Department of Special Education, and Lloyd Lovell, Department of Educational Psychology, both at the College of Education, University of Oregon; Dr. Adrian de Groot, Department of Psychology, University of Amsterdam, Holland, on sabbatical leave during 1967-68 at the Oregon Research Institute, Eugene, Oregon; and Dr. Gary Neilson, Psychological Associate, Psychology Department, Fairview State Hospital and Training Center, Salem, Oregon. Thanks are also extended to Drs. James Kiefert and Hugh Taylor, both of the Department of Education, College of Education, Washington State University, for advice received regarding the statistical procedures employed in the analysis of data.

Data collection for the study was made possible through the cooperation of Mrs. Jan Walter, Psychological Associate, Psychology Department, Fairview State Hospital and Training Center; Mr. Ed Skinski, Principal of the School at Fairview State Hospital and Training Center; Dr. Weston Niemela, Director of Special Education, Salem, Oregon; Mr. Rufus Franz, Principal of John F. Kennedy Junior High School, Eugene, Oregon; and the many officers, teachers, and secretaries of the various participating organizations.

To John C. Rolland and Jack W. Martin, who served as experimenters in the data collection phase of the experiment, a heartfelt "thank you" is extended.

A special indebtedness is to Cherie Warner, who served as principal typist, proofreader, critic, and personal source of encouragement during the project tenure.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ii

LIST OF TABLES. iii

Chapter

I. THE PROBLEM. 1

The Specific Problem 3
The Hypotheses 4
Crucial Terms. 5

II. REVIEW OF LITERATURE 7

The Relationship Between Goal Setting and Performance in Non-Mentally Retarded Subjects. 7
The Relationship Between Goal Setting and Performance in Mentally Retarded Subjects. 13
Summary. 16

III. PROCEDURES 22

Pilot Study. 22
Main Study 28

IV. RESULTS. 33

Testing the Major Hypotheses 33
Testing of Auxiliary Hypotheses. 36
Discussion of Findings 39

V. SUMMARY AND IMPLICATIONS 42

Problems 42
Procedures 43
Main Study 45
Limitations of the Study 48
Implications for Classroom Teaching. 49
Implications for Future Research 50

BIBLIOGRAPHY. 51

APPENDICES:

A. List of School, CA, MA, and IQ, for Non-institutionalized and Institutionalized Pilot Test Subjects. 54

B. Copy of Booklet Used in Pilot Testing. 57

TABLE OF CONTENTS - Continued

Page

C. List of Item Difficulties for Institutionalized and Non-Institutionalized Pilot Subjects on Each Spelling and Arithmetic Item 61

D. Mean Item Difficulties, Scores on Pilot Tests, CA, MA, and IQ for Each Main Study Subject (Grouped According to Treatment Condition) 63

E. Sample of Stimulus Cards (Items) Presented to Main Study Subjects. 66

F. Sample of Record Form Used by Experimenters to Record Predicted Scores and Attained Scores for Individual Main Study Subjects. 67

G. Raw Scores for Each Main Study Subject on Each Trial. . . . 68

LIST OF FIGURES AND TABLES

		Page
Table 2.1	Summary of Research Concerned with the Relationship Between Goal Setting and Performance in Non-Mentally Retarded and Retarded Subjects.	17
Table 3.1	Means and Standard Deviations on CA, MA, and IQ for Non-Institutionalized, Institutionalized, and Total Main Study Subjects	29
Figure 4.1	Schematic Representation of 2 X 2 X 2 Analysis of Variance Design Used in the Present Study	34
Table 4.1	Summary Table of Analysis of Variance Performed on Criterion (loth Trial) Scores.	35
Figure 4.2	Arithmetic Mean of Each of Eight Treatment Groups on Criterion (10th Trial) Scores. (N=10 for Each Group).	37
Table 4.2	Duncan's New Multiple Range Test Applied to Criterion Score Means for 8 Treatment Groups.	38

Chapter 1

THE PROBLEM

The professional literature related to studies of "level of aspiration" (LA) is voluminous. Most investigations, however, have been conducted under rather synthetic experimental conditions with little attempt being made to relate the findings to "real life" situations. As Fryer (1964) has observed, "...most of the studies tend to be laboratory rather than applied research, this constituting one of the limitations in the research literature on level of aspiration" (p. 3).

Because of the nature of the experimental tasks used and the conditions under which they have been employed, there are relatively few research findings with respect to goal setting behavior (the specifying of a "level of aspiration") which can be directly applied to an educational setting. Following the experimental technique developed by Frank (1935), common practice among most investigators has been to (1) present each subject with a task, (2) ask him to indicate how "well" he intends to perform the task, (3) observe as he performs the task, (4) report to him how "well" he actually performed, and (5) ask him to indicate how "well" he intends to perform on the next attempt, etc. With this technique "level of aspiration" is operationally defined as "...the level of future performance in a familiar task which an individual, knowing his level of past performance in that task, explicitly undertakes to reach" (Frank, 1935, p. 119).

Almost without exception, "...studies have employed level of aspiration as a dependent, rather than an independent variable..." (Fryer, 1964, p. 3) in an attempt to analyze and describe the phenomenon as it is affected by factors such as (1) amount of success and/or failure, (2) task difficulty, (3) magnitude of reward, (4) competition, (5) age and sex of subjects, (6) achievement need, and many others. Bayton (1943) has stated that "The experimental work on the level of aspiration has been characterized for the most part by concern with the analysis of the sources, tension-systems, or needs from which the level of aspiration arises" (p. 2) and has stressed the need for research dealing with the relationship between LA and actual performance. As Bayton (1943) has observed,

> ...the LA situation insofar as the individual is concerned does not terminate with the expression of his aspiration but continues into subsequent activities. If the needs described above have deeper significance than merely being determinants of aspiration, it would be expected that they would exert their influence upon behavior following the statement of aspiration. These

> needs...are not satisfied merely by the statement of the aspiration, but only by what occurs after that (p. 2).

Kausler (1959) has likewise observed that although "The research literature abounds with studies employing level of aspiration (LOA) as a dependent variable . . ." (p. 356),

> There is a paucity of information...concerning the relationship between LOA and subsequent level of performance (LOP) on the task. Consequently, little is known about the effects of LOA as a motivational independent variable. This is true in spite of the general acceptance of LOA as a motivational construct (p.356).

A fruitful approach for educational purposes would be to incorporate a principle such as goal setting behavior into an educational process and observe its effect upon behavior. In such a context the subject's goal setting behavior would be an independent variable, i.e., a "...variable that may be selected or changed by the experimenter ..." (Morgan, 1956, p. 633), an "antecedent condition" (Andreas, 1960, p.8). Peformance on the educational task would be the dependent variable, i.e., "...the variable that changes as a result of changes in the independent variable" (Morgan, 1956, p. 629).

The effect of individual goal setting upon task performance seems of particular interest in the training of educable mentally retarded adolescents. In recent years considerable effort has been made to provide more and better education for these students. Sniff (1963) has observed that

> ...emphasis on education for the educable mentally retarded has been placed to a greater extent on programs at the elementary school level than at the secondary school level. However, in very recent years more and more attention has been given to developing special education programs to adequately meet the needs of mentally retarded young adults at the secondary or high school level (p. v).

In general agreement with the above quotation is the statement by Robinson and Robinson (1965) that "Secondary school classes, for educable children at the junior and senior high school level, are becoming more prevalent as communities recognize how significant a contribution the retarded individual can make if he is properly prepared to assume his place in society" (p. 473).

Efforts are constantly being made within the field of mental retardation to develop curricula and teaching methods which will elevate the retarded student to his maximum potential. According to Hutt and Gibby (1958), the

> ...modern philosophical approach to the education of the retarded child is based on the concept of the worthwhileness of the child, and a belief that the child can be a contributing member of society. It recognizes his dignity as a human being, but also the fact that the child lacks many skills and talents. It is the function of special education to discover the specific capacities of each child, and to help him develop them to the highest possible level (p. 275).

Hutt and Gibby (1958) have argued that this modern philosophy of education is "...more realistic as well as more humane..." (p. 275) than earlier philosophies but that "...there is still a large gap between the proposed objective and what is actually achieved in educational practice" (p. 275).

The Specific Problem

The purpose of the present research project is to investigate the effects of the setting of goals upon the educational performance of educable retarded adolescents. Armstrong (1947), Fryer (1965), Kausler (1959), and Lockette (1956), in studying non-retarded subjects, have concluded that goal setting behavior has a strong motivational property; each investigator found that task performance scores are increased more under conditions where subjects state a "level of aspiration" before a trial than under conditions where subjects do not. If goal setting behavior can be found to have a similar effect upon the performance of educable mentally retarded adolescents on an educational task, this information can be incorporated into the classroom instructional process and thus better enable intellectually handicapped students to develop to their maximum potential.

Since goal setting as a training procedure has been found to be greatly influenced by level of task difficulty (Fryer, 1964), the relationship of this variable to task performance was also systematically evaluated in the present study. Tasks were constructed in such a manner that one-half of the subjects performed a task which was more difficult than the task performed by the other one-half of the subjects (see Chapter 3 for a more precise definition of tasks).

Numerous research projects have demonstrated educational differences between institutionalized and public school retardates of the same general IQ level. Because such differences were known to exist, the present study employed an equal number of subjects from each sub-population so that the effects of goal setting upon performance could be analyzed for institutionalized and non-institutionalized subjects separately. It was felt by the investigator that findings obtained with public school EMR males could not logically by generalized to the institutional population. For this reason, an equal number of institutionalized subjects were selected to permit the generalization of finds to that population.

Specifically, the following questions were studied in the research project:

1. Is there a difference between the subsequent performance of subjects who state performance goals (receive information on past performance and predict future level of performance) and subjects who only receive information on past performance?
2. Does this difference (if any) maintain similarly for subjects who perform an "easy" task and subjects who perform a "hard" task?
3. Does this difference (if any) maintain similarly for institutionalized and non-institutionalized subjects?

The Hypotheses

The primary objective of the present research project was to investigate the effects of goal setting upon subsequent performance on an educational task. The major hypothesis was stated as follows:

> H_1: The mean performance score on the criterion measure (10th trial) for subjects who state goals will be greater than the corresponding mean performance score for subjects who only receive information on past performance.

In other words, the mean performance score of subjects who state goals was, when averaged across levels of the two other variables (institutionalization and task difficulty), expected to be significantly greater than the mean performance of subjects who did not state goals. This hypothesis is suggested by the findings of Armstrong (1947), Fryer (1965), Kausler (1959), and Lockette (1956) which will be reviewed in detail in Chapter 2. Although none of these studies employed mentally retarded subjects, it is logical to assume that similar results will be obtained with retarded as with normal subjects.

In addition to the major hypothesis several auxiliary hypotheses were tested:

H_2: Non-institutionalized, "Hard" Task, Goal Setting subjects will have a greater mean performance score on the criterion measure (10th trial) than the corresponding mean performance score of Non-institutionalized, "Hard" Task, Non-Goal Setting subjects.

H_3: Non-institutionalized, "Easy" Task, Goal Setting subjects and Non-institutionalized, "Easy" Task, Non-Goal Setting subjects will not differ with respect to mean performance on the criterion measure (10th trial).

H_4: Institutionalized, "Hard" Task, Goal Setting subjects will have a greater mean performance score on the criterion measure (10th trial) than the corresponding mean performance score of Institutionalized, "Hard" Task, Non-Goal Setting subjects.

H_5: Institutionalized, "Easy" Task, Goal Setting subjects and Institutionalized, "Easy" Task, Non-Goal Setting subjects will not differ with respect to mean performance on the criterion measure (10th trial).

Hypotheses 2 through 5 are suggested, in part, by the findings of Fryer (1964) which indicate that goal setting is superior to non-goal setting in producing greater scores when a difficult task is involved but not when a less difficult task is involved (see Chapter 2 for a more complete review of literature). Hypotheses 2 through 5 are also built upon the assumption that goal setting on "hard" tasks will produce greater criterion scores in both institutionalized and non-institutionalized subjects than will non-goal setting.

Crucial Terms

In the present study goal setting behavior is defined as the act, on the part of a subject, of verbally expressing a performance level which he "thinks he can" achieve on an anticipated task. Level of aspiration is defined in this study as the performance level which the subject declares, i.e., the "goal" he sets.

Educable mentally retarded (EMR) adolescent is defined in the present study as (1) a child between the chronological age of approximately 13 and 18 years who has, on the basis of Oregon state law (see Chapter 3), been certified as educable mentally retarded and placed in a special class for the educable mentally retarded in a public school system, or (2) a child between the chronological age of approximately 13 and 18 years who has been placed in an institution for the mentally retarded and has, on the basis of the evaluations of staff psychologists, school administrators, and teachers, been placed in classes for the educable mentally retarded.

Educational task, experimental task, and task are used interchangeably in the present report and refer to a set or unit of 10 problems presented to each subject individually. "Easy" tasks and "hard" tasks differ in terms of the difficulty of items (established from pilot test data) of which they were composed. A more precise definition of "easy" task and "hard" task will be given in Chapter 3.

Trial is defined as a single complete presentation of each task set. In all cases, a task set was composed of 10 items, each item printed on a single 3 x 5 inch white card. Each item was presented individually to the subject and the successive presentation of all ten items constituted a trial.

The forty institutionalized subjects within the main study are sometimes referred to as the institutionalized sub-sample while the forty public school subjects are called the non-institutionalized sub-sample or, in some cases, the public school sub-sample.

The author of this report is frequently referred to as the investigator. The term experimenters has reference to the persons actually gathering main study data in face-to-face communication with the subjects.

Chapter 2

REVIEW OF LITERATURE

To understand the purposes of the present investigation it is necessary to consider in some detail the past findings of research related to "level of aspiration". The present chapter will, therefore, deal with a review of literature under two general areas: (1) The relationship between goal setting and performance in non-mentally retarded subjects, and (2) The relationship between goal setting and performance in mentally retarded subjects.

The Relationship between Goal Setting and Performance in Non-mentally Retarded Subjects

Early researchers of "level of aspiration" seem to have assumed that goal setting and subsequent level of performance were related. May and Doab (1937), for example, stated that "motivation is a function of the discrepancies between level of aspiration and level of achievement" (Gardner, 1940, p. 66). Frank (1941) further suggested that "level of aspiration"

> ...may be used to improve performance by being placed far enough above actual performance to act as an incentive. Conversely, if a high level of aspiration injures performance by making the subject tense, he may try to improve his achievement by lowering it (p. 224).

Only a few research findings directly relate to the general assumption that goal setting has a facilitating effect upon subsequent performance. Several early studies obtained conflicting results. Filter (1927), for example, in an early study which antedates the introduction of the term "level of aspiration" into the literature (Dembo, 1931), had a number of young people of both sexes (562 school children in grades ranging from fourth to ninth and 154 college sophomores) estimate their performance on six different tasks:

(1) How many makes of automobiles they could name in 3 minutes;

(2) How many capitals of states in the U. S. they could name in 3 minutes;

(3) How many different magazines they could name in 3 minutes;

(4) How many of 16 nonsense syllables they could recognize 10 minutes after presentation;

(5) How many of 40 fractions they could express in per cents if given 3 minutes to work; and

(6) How many of 40 columns of figures they could correctly add in 3 minutes.

After all estimates had been made, each subject (S) was asked to perform the six tasks.

Although Filter's main concern was with "...conditions of correctness in estimates and constancy of over- or under-estimation in varying situations..." (Filter, 1927, p. 58), he does report Pearson Product Moment Coefficients of correlation between estimates and performance for the 154 college sophomores. These correlation coefficients were .32, .14, .24, .51, -.10, and .13 for each of the six tasks respectively. Filter also reports the results of a separate study where 48 college students predicted their performance on (1) a dart throwing task, and (2) a ball throwing task. Spearman Rank Difference Coefficients between estimates and performance were -.14 for dart throwing and .05 for ball throwing.

The findings of Filter (1927) suggest that there is no clear-cut relationship between predicted scores and obtained scores in the type of tasks employed.

Kneeland (1934) administered three tasks to each of 406 subjects (college and high school students, and department store workers involved in an executive training course). Tasks were: (1) the cancellation of numbers containing both a 6 and a 9 in a series of numbers, (2) a "scribbling task" in which the S made as many oscillations with a pencil while lowering it down the page as quickly as possible, and (3) a knot-tying task consisting of typing knots as quickly as possible in a piece of twine a yard in length. Each task was performed 4 times and before the second, third, and fourth trials the S was asked to predict his performance.

Kneeland (1934) concluded, among other things, that "While there was a positive correlation between pre-estimates of improvement and actual improvement, the coefficients were not sufficiently high to indicate an individual's improvement from his estimate of improvement before the trial" (p. 70).

Yacorzynski (1942) had 40 subjects (16 males and 24 females; average age 38.4 years, range 12 to 58 years) each perform four tasks: (1) a tapping test on which the S was instructed to make 2 dots with a pencil in each square on a page ruled into 10 by 15 squares, (2) a substitution test in which the S was asked to place different symbols under their appropriate numbers, (3) a bead stringing test in which the S placed alternating blue and yellow beads on a string, and (4) a word association test in which the S was asked to "...name different words" (p. 407). Four trials were given on each of the four tasks. After each trial the S was told his performance score and asked to state the score he expected to make on the next trial.

Although Yacorzynski's study was not undertaken primarily to determine the relationship between LA and performance, the experimenter did find that there was a tendency for performance to be higher for subjects having the highest aspirations. These results are difficult to interpret, however, since the relationship between performance and LA was confounded with other variables within the study.

Bayton (1943) tested 300 Negro women attending Virginia State College on each of two tasks: (1) an arithmetic task composed of problems such as the multiplication of 4 and 5 digit numbers, long division, and addition, and (2) a cancellation task. Subjects were matched for "...previous performance ability and subjective estimate of the performance..." (p. 9). Ss were asked, prior to the criterion trials, to state their Maximum, Actual, and Least levels of aspiration. Analysis of the data led Bayton to conclude that for the ego involved task (arithmetic) there was "...a reliable tendency for those subjects with higher Actual levels of aspiration to follow them with better performances" (p. 20).

The studies of Filter (1947), Kneeland (1934), Yacorzynski (1942), and Bayton (1943) have several similarities worth noting. First, all are concerned with the relationship between stated goals and subsequent performance (although, it should be noted that Kneeland (1934) was concerned with predictions of increases in performance). Second, all of the studies attempt to find a relationship between the magnitude of LA and the magnitude of performance, i.e., all subjects stated a level of aspiration and the magnitude of that level and not its presence or absence is of concern. A third similarity among the studies is the nature of the tasks employed. Although the tasks differed considerably as to the exact muscle responses required by the S (stringing beads, tapping a pencil, crossing out letters, etc.) all tasks were essentially of a speeded nature; even the arithmetic problems used by Bayton (1943) were selected to that the S could work the problem if she reached it during the time limit imposed.

In addition to the several similarities, the previously cited studies are different in at least one important way; the findings of Filter (1927) and Kneeland (1934) suggest a low relationship between estimates and performance while the results of the work of Yacorzynski (1942) and Bayton (1943) support the idea that higher LA leads to higher performance. Such differences may, perhaps, be attributed to differences in subjects, differences in tasks employed, and/or to the different methods used for analysing the data; the exact reason for such obtained differences will, of course, remain a matter of conjecture since there is no reliable way to identify the source of differences. The statement by Bayton (1943), made over two decades ago, seems as true today as when it was written.

> When we review the literature on the level of aspiration, we find that despite the fact that the term implies a subjectively determined goal toward which one is striving, there are few experiments which have tested the hypothesis that level of performance is in any way dependent upon the height of preceding aspiration (p. 2).

A second group of investigators (Armstrong, 1947; Fryer, 1964; Kausler, 1959; and Lockette, 1956) have attempted to determine the influence of LA upon subsequent performance through procedures quite different than those of the previously discussed researchers. This latter group of investigators has been concerned more with the presence or absence of a stated LA (i.e., with the expressing of a goal) than with the magnitude of the LA. Common to all these studies has been the procedure of asking one group of subjects to estimate future performance while a second group of subjects made no such predictions.

According to Fryer (1964), Armstrong (1947) was the first experimenter to employ the above research procedure. She compared the performance of subjects under "expressed" and "non-expressed" levels of aspiration against a baseline of knowledge of results alone. The task involved was the Minnesota Rate of Manipulation Test. Although her results revealed that goal setting was superior to knowledge of results alone in producing high performance scores, the findings are open to criticism since the knowledge of results, (i.e., the feedback information furnished subjects concerning their past performance) was ficticious. As Fryer (1964) has observed:

> Although a uniform series of fictitious performance scores does have the advantage of standardizing the reported level of performance and the shape of the learning curve for all subjects, there is an explicit disadvantage entailed in this procedure. Under such arrangements, certain tasks are of questionable appropriateness, since the subject's true level of performance must not be obvious to him. There is always the possibility that a subject may sense changes in his performance which appear to be contradicted by the scores reported to him, resulting in skepticism concerning the reported information.... Certainly it would seem that reporting the subject's actual performance to him and asking him to set his aspiration level within that framework would make for a much more realistic situation (p. 15).

Lockette (1956) had "...untrained junior and senior high school subjects..." (p. 284) plane a piece of wood to specified pre-set dimensions. Each S performed the task 6 times. Part of the subjects were required to set performance goals and part of the subjects were not. Instructions varied so that part of his goal setting subjects set "realistic" levels of aspiration and others set "unrealistic" levels. Results indicated, among other things, that "...Goal-setting, either realistic or unrealistic, is superior to no goal-setting at all" (p. 284).

At least two serious weaknesses are apparent in Lockette's study. First, the control subjects were not told how the task was to be graded; second, and more serious, the control group received no feedback from the experimenter concerning their performance as the tasks progressed. As Fryer (1964) has observed, "His [Lockette's] study, in fact, does not indicate whether sheer knowledge of results or the goal-setting aspect of level of aspiration, or both, accounted for the enhanced performance" (p. 14).

Kausler (1959) had 118 male and female undergraduates at the University of Arkansas complete a 25 item arithmetic (practice) test and a 50 item arithmetic test. Approximately one-third of the students did not set goals, one-third set an individual goal, and the remaining one-third set an individual goal after being told what a minimum acceptable level of performance would be. All subjects worked for 3 minutes on the practice test and 6 minutes on the main (50 item) test. The expressing of a LA took place after the practice test and constituted the subject's writing, on the test booklet, the number of problems he or she "hoped" to be able to answer during the 50 item test.

In analysing his data, Kausler used the practice test results for each student as "...a control variable for differences in ability on the arithmetic task" (p. 348). Using an analysis of covariance design, Kausler obtained results indicating that the mean perfromance of the subjects who set goals was significantly greater than the mean performance of subjects who did not. He concluded that "...expressing an aspiration level served to increase performance level on the subsequent task" (p. 350).

The research of Kausler has at least two weaknesses which should be briefly considered. First, it differs from traditional LA studies in that the task was presented only once to each S. The S, therefore, had no opportunity to determine his performance on the actual task before setting a goal. The 25 item practice test, although similar in nature, provided the subject, at best, with only a general idea of his performance ability. The second weakness, related closely to the first, is the fact that S received no clear knowledge of his performance on the 25 item exam before attempting the main test. The practice exam was not corrected before the main test was taken so the S did not clearly know how many of the practice items he had correctly answered. The specifying of goals, then, was not made on the basis of clear, accurate knowledge regarding past performance.

Fryer (1964) had one hundred male freshmen and sophomores from Boston University learn to accurately receive International Morse Code. Prior to the main study, the difficulty level of various letters was determined and instructional procedures were standardized. In the main study, half of the subjects received "high difficulty code characters" and half received "low difficulty code characters". Each S had 15 trials with the summation of scores on trials 6 through 15 serving as the criterion measure. In addition to difficulty level, Fryer also studied the effects of goal setting vs. non goal setting, as well as two other variables which are not of direct concern to the present review.

Analysis of covariance, employing the Ss initial performance as a control variable, revealed a significantly larger mean performance for Ss who set goals than for the Ss who did not. Further analysis, however, revealed that this relationship was true only for Ss performing the difficult task; the group means of goal setters and non goal setters were not significantly different on the easy task considered alone.

Fryer's study possesses several strengths not seen in the previously considered projects. First, the subjects received accurate feedback information at the end of each trial concerning their previous performance thus providing a realistic basis for future estimating. Differences in performance between groups could, therefore, be considered in terms of the variable "knowledge of results plus goal setting vs. knowledge of results alone".

A second strength of Fryer's study is that two different difficulty levels were studied within the same general project.

A third strength of Fryer's study evolves from the fact that it suggests some practical application; the teaching of Morse Code for military or civilian use may be facilitated by incorporating Fryer's techniques. The previously used tasks (manipulation tasks, planing a piece of wood, special arithmetic items, etc.) have more limited possibilities for application.

Although the studies of Armstrong (1947), Fryer (1964), Kausler (1959), and Lockette (1956) reflect major differences with regard to the subjects studied, the kinds of tasks employed, and the conditions under which a LA is expressed, they possess one major similarity--each study obtained evidence that goal setting subjects, as a group, attain a higher level of performance than a similar group of subjects who do not set goals.

It should be kept in mind that none of the above mentioned studies have involved mentally retarded persons. Results, therefore, cannot be justifiably generalized to include such individuals. The question as to whether or not retarded persons will attain a higher level of performance on a given task when asked to set a goal can only be answered on the basis of empirical evidence gathered specifically from retarded subjects.

The Relationship between Goal Setting and Performance in Mentally Retarded Subjects

As with studies of "level of aspiration" in general, the vast majority of investigations of goal setting among mentally retarded subjects have dealt with LA as a dependent variable. An excellent review of research related to the LA phenomenon in retarded persons has been made by Cromwell (1963). Though of considerable interest to the student of mental retardation, the majority of studies are not related to the presently reported research project and, therefore, will not be reiterated here. Two studies, however, have focused upon the relationship between LA and performance; these studies are presented in some detail in the following paragraphs.

Eyman (1964) tested 45 male subjects 21 years of age and older on a rail-walking apparatus. Fifteen subjects were placed in each of the following IQ groupings: 50-59, 60-69, 70-80.

Each S walked the rail 7 times, beginning each time at the wide end (4 inches) and proceeding toward the narrow end (½ inch in width). Before each trial (except the first) the S was asked to indicate: (1) How far he would have to go in order to get a "good score", and (2) What a "poor score" for him would be. Half the distance between these points was defined as the S's "level of aspiration". The performance score for each trial was defined as the last place S stepped before stepping off the rail.

Results showed a Pearson Product Moment Correlation Coefficient of .660 between performance and "level of aspiration'. This relationship was linear and was interpreted by the investigators as "...supporting evidence indicating that the arithmetic mean of performance based on previous trials was acting as the adaptation level for subsequent aspirations" (p. 749). Results also indicated that subjects with higher IQs set their LA more realistically (closer to performance scores) than did Ss with lower IQs.

Subotnik (1967) attempted to relate level of aspiration and learning in 28 institutionalized educable mentally retarded boys (CA range = 9-6 to 11-8, MA range = 5-4 to 8-5, IQ range = 47 to 92). The rather complex data collection procedures are described by Subotnik (1967) as follows:

> Materials for the level of aspiration task were simple pictures--a puppy, a rabbit, and a Santa Claus--taken from a reading readiness workbook, presented in randomized order (except that no two consecutive pictures were the same) and aligned along the edge of a table. The child was given a set of poker chips in three colors and shown on a model provided through the task that each picture was assigned a certain color. He was told to place as many chips under the

pictures as he could in the given time, beginning at the left. The amount of time was manipulated by *E*. After each of 11 trials, *E* moved a marker with a red star to the last picture reached. The boy was asked, "How far do you think you will get next time?" and placed a marker with a gold star at the picture he expected to reach on the next trial. The pictures were numbered consecutively to provide a score for *E* to record. *S*'s errors were ignored, except that *E* called attention to them in the first trial or so to induce a set for accuracy.

On the first trial *S* was allowed to reach the midpoint of the array of pictures, the 14th. His succeeding scores were 16, 18, 20, 17, 15, 13, 13, 10, 13, and 16. *E* attempted to accentuate feelings of success and failure by comments of approval after a rise in performance (trials 2, 3, 4, 10, 11) and by deprecatory comments after a decline in or failure to improve performance (trials 5 through 9) regardless of whether *S* attained his own prediction.

The learning task required *S* to learn an arbitrary association of seven colors with seven simple pictures (knife, horse, boat, bucket, radio, baby, cake) taken from a reading readiness workbook. A picture was presented, mounted on the front of a folded card; *S* pointed to one of seven colored squares on a key before him, after which the card was opened to expose a colored square, confirming or disconfirming *S*'s response. After the preliminary exposure the series of pictures was presented in varied order 20 times or until *S* attained five consecutive perfect series. The score consisted of the total number of errors made.

Two teachers, who knew all *Ss*, served as raters of classroom learning performance and of emotional disturbance. A "reading readiness" classroom teacher (Judge A) and a handwork teacher (Judge B) were given the *Ss*' names on small individual cards and asked independently to arrange them in order (a) "from the best learner--the one who masters his material well and retains it well--to the poorest learner" and (b) "from most disturbed emotionally--not necessarily most troublesome--to least disturbed emotionally."

Ss were classified: (a) in two equal groups, wide or narrow, in range of discrepancy scores and (b) by tendency to maintain a low positive discrepancy throughout the trials (10 *Ss*) vs. other tendencies (18 *Ss*), such as high positive, negative or mixtures, which were combined rather than treated separately because of the small number of *Ss*. The low positive group was required to have at least eight discrepancies on the eleven trials in the range 1 to 5 and a total of all

> trials in the range 18 to 40. These criteria were somewhat arbitrary and determined upon inspection of the data.
>
> The groups then were compared as to performance on the learning task and the teachers' ranking on classroom performance and emotional disturbance (pp. 768-769).

Results of a non-parametric Mann-Whitney comparison of ranks indicated that Ss with a low positive discrepancy on the level of aspiration task "...did more poorly on the learning task than those with other goal setting patterns" (p. 769) although the groups did not differ as to IQ, MA, and CA (P = .001 level). In other words, those subjects who kept their level of aspiration slightly higher than their past performance throughout the series of trials on the LA task had a lower performance, as a group, on the learning task than subjects with other goal setting patterns, i.e., subjects who set LA much higher than past performance, lower than past performance, and/or the same as past performance. The low positive discrepancy group was not found to differ from others in classroom performance (as judged by teachers) and was found to be superior, according to teacher ratings, in emotional adjustment. Subotnik (1967) also reports that "Ss with the narrower range of goal discrepancy scores did not differ significantly on the learning task from those with a wider range" (p. 769).

Unfortunately, the author did not report the exact discrepancy scores of the group who had other than "low positive discrepancies". If these scores were primarily "high positive" in nature then, since the actual obtained score on the LA task was controlled by E and, therefore, was equal for all Ss, we could explain the obtained results in terms of goal setting, i.e., Ss who achieved the highest learning scores set the highest LA on the initial task. Subotnik (1967) suggests that the Ss who achieved higher learning scores (and who were more emotionally disturbed) may have experienced "...a more intense drive state (i.e., anxiety)" (p. 769).

Another possible explanation for Subotnik's unexpected findings, and one which he did not discuss himself, is that the two tasks employed may be so different that LA predictions on one should not logically be expected to relate to performance on another. Unlike the studies of Armstrong (1947), Fryer (1964), and others previously discussed in this chapter, the research of Subotnik did not permit Ss to set goals on the performance task itself. There is not sufficient evidence to conclude that LA on one task has a high relationship to aspiration on another. Therefore, Subotnik's findings may be a result of differences between tasks.

The studies of Eyman (1964) and Subotnik (1967) have few similarities other than the fact that both employed mentally retarded subjects (even then, the chronological ages of the two samples differ greatly). Eyman (1964) was concerned with the relationship between LA and performance on a motor-type skill. Subotnik (1967) was concerned with the LA discrepancy scores on a sorting task as they related to learning on a school-like task. The general finding of Eyman (1964), that there was a significant relationship between LA and performance, is in general agreement with the studies of Yacorzynski (1942) and Bayton (1943) on non-retarded subjects. The findings of Subotnik (1967) are not easy to explain and do not seem to fit well with any of the previously reviewed studies.

Summary

Table 2.1 provides a very brief summary of the research studies reviewed in the preceding sections of this Chapter. Studies are listed in chronological order according to publication date. In cases where the major foci of the investigation were not of concern to the present review, i.e., where the relationship between LA and performance was of secondary concern to the original researcher, double parentheses have been placed around the statement in the "Study Foci" column of Table 2.1.

In conclusion, it is probably most accurate to say that studies designed to investigate the relationship between the magnitude of LA and the magnitude of subsequent performance scores have obtained inconclusive results. On the other hand, studies comparing the mean performance of subjects who set goals with the mean performance of subjects who did not have, without exception, found overall superior performance on the part of goal setting subjects. None of these later studies, however, has involved mentally retarded subjects. The present project is an attempt to investigate the effects of goal setting upon the preformance of male educable retarded adolescents on spelling tasks of differing difficulty levels.

TABLE 2.1

Summary of Research Concerned with the Relationship Between Goal Setting and Performance in Non-Mentally Retarded and Retarded Subjects

---

Filter, R. O. Estimates of the amount of work one can do. J. appl. Psychol., 1927, 11, 58-67.

Study Foci: (("...conditions of correctness in estimates and constancy of over- or under-estimation in varying situations (p. 58).))

Sample: 562 school children in grades 4 to 9; 154 college sophomores (Male and Female Ss). 48 additional college students

Task Description: Six tasks were used: (1) Number of makes of automobiles named in 3 minutes. (2) Number of capitals of states named in 3 minutes. (3) Number of different magazines named in 3 minutes. (4) Number of nonsense syllables recognized 10 minutes after presentation. (5) Number of fractions expressed in per cents in a 3 minute period. (6) Number of columns of figures added correctly in 3 minute period. 48 college students also predicted scores on (1) dart throwing, and (2) ball throwing tasks.

Administration of Task: Ss were asked to estimate their performance on each task. After estimates were given, each S performed each task described.

Results: Pearson product moment r's between predicted performance and actual performance were .32, .14, .24, .51, -.10, and .13 for each of the six tasks respectively for 154 college sophomores. For 48 additional college students Spearman Rank Coefficients of -.14 (dart throwing) and .05 (ball throwing) were found between estimates and performance.

Kneeland, N. Self-estimates of improvement in repeated tasks. Arch. Psychol., 1934, 163, 1-75.

Study Foci: Relationship between estimates of improvement and actual performance on three (3) tasks.

Sample: 406 subjects (male and female) from college, high school, and dept. store executive training class.

Task Description: Three tasks were employed: (1) cancellation of numbers containing both 9 and 6, (2) "scribble test", (3) knot-tying.

Table 2.1 (continued)

Kneeland (continued)

Administration of Task: Ss were presented each task and then, prior to the 2nd, 3rd, and 4th presentations were asked to predict future performance levels.

Results: Low positive Pearson Product Moment r's between "...pre-estimates of improvement and actual improvement..."

Yacorzynski, G. K. Degree of effort. III. Relationship to the level of aspiration. J. exp. Psychol., 1942, 30, 407-413.

Study Foci: ((Relationship between the degree of effort expended on a task and the LA))

Sample: 16 males and 24 females; average CA of 38.4; range of CA from 12 to 58 years.

Task Description: Four tasks were used: (1) tapping test, (2) substitution test, (3) bead stringing test, (4) word association test.

Administration of Task: Four trials were given to each S on each task. After each trial S was told his performance score and asked to predict the score he expected to make on the next trial.

Results: "Tendency" for performance to be higher for Ss having highest aspirations. Results confounded with other variables.

Bayton, J. S. Interrelations between levels of aspiration, performance, and estimates of past performance. J. exp. Psychol., 1943, 33, 1-21.

Study Foci: Relationship between expressed levels, i.e., Maximum, Actual, and Least, of aspiration and subsequent performance.

Sample: 300 Negro college women.

Task Description: Two tasks were employed: (1) arithmetic task, (2) cancellation task.

Administration of Task: Ss asked to state Maximum, Actual, and Least levels of aspiration before performing task.

Table 2.1 (continued)

Bayton (continued)

Results: On arithmetic task there was "...a reliable tendency for those subjects with higher Actual levels of aspiration to follow them with better performance" (p. 20).

Armstrong, D. D. Performance as a function of expressed and non-expressed levels of aspiration. Unpub. Master's thesis, Howard University, 1947. As reviewed by Fryer, F. W. in An evaluation of level of aspiration as a training procedure. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Study Foci: Comparison of performances under conditions of "expressed" and "unexpressed" levels of aspiration with performance under "knowledge of results alone".

Sample: ?

Task Description: Minnesota Rate of Manipulation Test.

Administration of Task: Some Ss expressed goals prior to task performance and others did not. Ss were given "feedback" on a predetermined schedule independent of actual performance.

Results: "Goal setting" superior to "knowledge of results alone" in in terms of performance scores.

Lockette, R. E. The effects of level of aspiration upon the learning of skills. Dissertation Abst., 1956, 16, 284 (Abstract).

Study Foci: Comparison of performances under conditions of "realistic" and "unrealistic" goal setting with performance under non-goal setting conditions.

Sample: "...untrained junior and senior high school subjects..." (p. 284).

Task Description: Planing of a piece of wood to pre-set dimensions.

Administration of Task: Students given instructions for planing wood. Some Ss were asked to set goals and some were not. Task was repeated 5 times, i.e., for a total of 6 trials.

Results: "Goal setting...superior to no goal setting at all" (p. 284); i.e., performance scores were higher under goal setting conditions than non-goal setting conditions.

Table 2.1 (continued)

Kausler, D. H. Aspiration level as a determinant of performance. J. Pers., 1959, 27, 356-361.

Study Foci: Comparison of performance under "goal setting" conditions with performance where such setting of goals is not required.

Sample: 118 male and female undergraduates.

Task Description: 50 item arithmetic test.

Administration of Task: Ss were first given a 25 item arithmetic test (3 minutes) following which part of them were asked to predict performance of a 50 item arithmetic test (6 minutes).

Results: Ss who specified goals did better on 50 item test even when "ability" was controlled through analysis of covariance design.

Fryer, F. W. An evaluation of level of aspiration as a training procedure. Englewood Cliffs, N. J.: Prentice-Hall, 1964.

Study Foci: The effect of (1) goal setting, (2) difficulty level, (3) method of instruction and (4) conditions under which LA was expressed upon subsequent performance.

Sample: 100 male frosh and sophomore college students.

Task Description: Decoding words from Morse Code.

Administration of Task: Each subject was given 15 trials and after each trial S predicted his next performance.

Results: Main effects in the analysis of covariance design were found significant for goal setting and difficulty level. Method of instruction and conditions for expressing LA were not significant. Group means differed between goal setting and non-goal setting for difficult tasks but not for easy. Goal setting superior.

Eyman, R. K. Covariation of level of aspiration and adaptation level with other characteristics. Amer. J. Ment. Defic., 1964, 68, 741-749.

Study Foci: Relationship between LA and subsequent performance.

Sample: 45 male MR Ss 21 years of age or older.

Table 2.1 (continued)

Eyman (continued)

Task Description: Rail walking.

Administration of Task: Ss asked to specify before each trial (except #1) (1) How far along the rail they would have to go to get a "good" score, (2) What a "poor" score for them would be. LA defined as midway between 1 and 2 above. A total of 7 trials were conducted.

Results: Significant r (.660) between LA and performance.

Subotnik, L. Level of aspiration, emotional disturbance, and learning in institutionalized educable mentally retarded boys. Am. J. Ment. Defic., 1967, 71, 767-771.

Study Foci: Relationship between LA goal discrepancy, learning, and maladjustment.

Sample: 28 institutionalized EMR males (CA range of 9-6 to 11-8; MA range of 5-4 to 8-5; IQ range of 47 to 92).

Task Description: LA task consisted of placing different colored poker chips on different pictures. Learning task required S to learn an association of 7 colors with 7 pictures.

Administration of Task: Ss given LA task under conditions where performance was controlled by E. Ss then given learning task.

Results: Ss with low positive discrepancy scores on LA task obtained lower learning task scores than Ss who had other discrepancy score patterns.

Chapter 3

PROCEDURES

Pilot Study

The purpose of the pilot study was twofold: (1) the development of appropriate educational tasks for use in later testing, and (2) the standardization of instructions and experimental procedures.

Development of an Educational Task

Initial Attempt. It was originally planned that 5 institutionalized and 5 special class subjects would be selected at random from their respective populations and tested on an individual basis. Three tasks were selected by the investigator: (1) a paired-associate learning task developed by Drew, Prehm, and Logan (1967), (2) a spelling task constructed by the investigator using words taken from an instruction manual for educable mentally retarded students (Beloit, Wisconsin Public Schools, 1965), and (3) an arithmetic test constructed by the investigator using simple, single and double digit addition, subtraction, multiplication and division problems. Within each task a single problem was presented on one 4 x 6 inch white card. Tasks ranged from 14 cards (paired-associate learning) to 36 cards (spelling) in length.

Testing procedure was for the subject to sit at a desk with the investigator seated beside him. Instructions differed slightly from task to task but in all cases the subject was informed that he would be presented with a problem and was to give a correct answer within 15 seconds.

As testing proceeded, it soon became apparent that wide individual differences existed between subjects. The first subject, for example, was able to get only 9 of the 24 arithmetic problems correct on the first trial and only 10 on the third trial. S #2 answered 21 of 24 correct on the first trial and 24 of 24 on the second. In spelling, S #1 got only 1 of 10 correct on the initial trial and 3 of 10 correct on the fourth trial; S #2 missed only 1 of 36 spelling words on the first trial and none on the second trial. With regard to paired-associate learning, S #2 learned 14 high association pairs to criterion (all correct) in 3 trials; S #1 could not read the words well enough to begin the task.

The investigator was informed by Fairview personnel that such wide individual differences could be expected throughout the entire population being investigated. On the basis of their recommendation and the obvious evidence that wide differences existed, it was concluded that the testing of only 5 institutionalized and 5 non-institutionalized subjects would not constitute a sufficient pilot study. The initial plan for collecting pilot data was, therefore, abandoned.

Group Testing. The initial plan for collecting pilot data suffered from at least two weaknesses. First, it involved only a few subjects who, because of the marked heterogeneity of the population, could not be taken as being representative; second, only a limited number of items could be used since individual testing demanded a considerable amount of time. To alleviate these two major problems it was decided that a group administered paper and pencil test should be given. This would permit the gathering of initial data from a large number of pilot subjects in a minimum of time and would enable the "trying out" of a large number of items of varying difficulty. The results of this paper and pencil test not only would enable the investigator to determine the empirical difficulty of each item but would also yield information regarding the ability of each potential subject for the main study.

Two general types of problems--arithmetic and spelling--were judged by the writer as appropriate for investigation as possible problems for use in the main test to follow. The types of problems were required to meet the following criteria: (1) the problems must be educational in nature, i.e., they must deal with subject matter which is generally considered an appropriate part of the classroom educational process for the subjects involved, (2) the problems must be of such a nature that "learning" can occur within the testing sessions, i.e., performance scores can increase as a result of information presented as part of the testing process, and (3) the problems must be of varying difficulties (as determined by the percentage of pilot subjects correctly answering each item).

Results of the recently administered Wide Range Achievement Test (Jastak and Bijou, 1946) and the findings of a past programmed reading experiment indicated that the population from which institutionalized subjects were to be selected suffered, in general, a severe reading deficit. For this reason the investigator judged the paired-associate task inappropriate and elected to abandon it in favor of the arithmetic and spelling tasks.

Subjects for the group pilot test were 93 educable mentally retarded male adolescents between approximately 13 and 18 years of age. Forty-eight (48) of the subjects were from Fairview State Hospital and Training Center in Salem, Oregon; twenty-one (21) subjects were from special education classes at John F. Kennedy Jr. High School in Eugene, Oregon; twelve (12) subjects were from the Special Education Center, Salem, Oregon; and the remaining twelve (12) were from special education classes at North Salem High School, Salem, Oregon. Appendix A reports the chronological age, mental age, and IQ score for each of the pilot test subjects for whom the information was available. Means and standard deviations are also given for institutionalized, non-institutionalized, and total sample on each of the above measures.

The 48 institutionalized subjects represent the entire population of available educable retarded males between approximately 13 and 18 years of age at Fairview State Hospital and Training Center. The 21 subjects from Kennedy Junior High were all of the 25 EMR males at that school who were present at school during the test administration. The 12 subjects at the Special Education Center were all of the original 13 who were present at school when the test was given; at North Salem High School 12 out of a possible 19 were present for testing.

Mentally retarded children in the public school system of the state of Oregon must be individually certified according to state law. By legal definition:

> "Mentally retarded children" means children between the ages of 6 and 21 who because of well-established retarded intellectual development are incapable of receiving a common school education through regular classroom instruction but whose intellectual ability would indicate a possible scholastic attainment of third-grade level with the benefit of special instructional methods; who are competent in all aspects of the school environment except the academic (Chapter 541, Oregon Laws 1961).

Certification is obtained from the State Superintendent of Public Instruction and "Each application for certification of a child as mentally retarded draws upon three sources of information; developmental and school history, individual intelligence testing, and a statement of a physician following physical examination" (Purdom, 1963, p. 2). While educable mentally retarded children are commonly defined as having IQ scores between 50 and 75 (Kirk, 1961) the laws of Oregon do not specify IQ restrictions and it is not uncommon to find special class students with IQ scores reaching into the low and mid 80s. In the present study, for example, the mean IQ for the public school pilot subjects was 80.00 with a range from 112 to 56. As is evident from an examination of Appendix A, the subjects from Kennedy Junior High School had a tendency to score higher than would be expected for students in classes for the educable mentally retarded. Possible reasons for these unusually high scores are discussed in Appendix A.

Educable mentally retarded persons living within an institution are not necessarily certified by the State Superintendent of Public Instruction although many of them may have been so certified as special education students before being institutionalized. An institutionalized person at Fairview State Hospital and Training Center is placed in a special class for the educable retarded when, on the basis of IQ scores, past history of schooling, and present teacher and principal opinion it is judged that he can profit from such a program.

At Fairview State Hospital and Training Center male patients who have a history of "delinquency" and are judged as behavioral deviates by institutional personnel are housed in one specific cottage. In selecting subjects an effort was made to avoid drawing patients from this cottage. While it is true that some of these persons function at the educable mentally retarded level in an academic setting, they do not represent the typical institutionalized retardate (i.e., they have behavior problems in additional to those common among retarded persons) and were, therefore, not included in the sample.

Provision had been made, in advance of testing, for removing those Ss who had a physical impairment which was judged by the investigator to render task performance impossible. However, no such S appeared for testing.

A complete list of arithmetic items and spelling words presented to the pilot study subjects is presented in Appendix B. The proportion of correct responses to each item (item difficulty) is presented in Appendix C.

Arithmetic items for the test were constructed by the investigator and were patterned after those suggested in available curriculum guides (Department of Education, State of Alaska, 1967; The Board of Public Education, Pittsburgh, Pennsylvania, 1963a, 1963b, 1963c) as appropriate for educable mentally retarded students in junior and senior high school. Spelling items were selected at random from texts prepared for use with non-retarded students in grades 2 through 7 (Horrocks, Evans, & Staiger, 1965; Horrocks, Linsenmeier, and Staiger, 1965; Horrocks, Sachett, & Staiger, 1965; Horrocks & Staiger, 1965a, 1965b; Glim & Manchester 1967a, 1967b, 1967c, 1967d, 1967e, 1967f). It was felt by the investigator that these words would provide a satisfactory range of difficulty for students who, by definition, are expected, at best, to do academic work somewhere between the second and sixth grade level (see Kirk, 1961, p. 110). Words selected from one language arts curriculum (Beloit Public Schools, 1965) were also included and, to assure that a wide range of item difficulties was attained, several alphabetic characters selected from the Wide Range Achievement Test (Jastak and Bijou, 1946) were placed at the front of the spelling exam.

Arithmetic problems were presented on a printed page and subjects were given a maximum of 40 minutes to complete them. Spelling words were read aloud twice by the experimenter and subjects were given a maximum of 15 seconds following the second reading to write the word in the appropriate blank on the test form. First and second presentation of spelling words were 2 seconds apart. (See Appendix B for a copy of the test booklet.)

Conrad (1951) has explained that "Since one of the purposes of a tryout is to gather accurate data about each individual item, it is extremely important that the time limits be so generous as to permit all, or nearly all, of the examinees to attempt to answer every item on which tryout data are desired" (p. 258).

The time limits permitted for the arithmetic and spelling sections of the pilot test appeared to the investigator as very adequate. For computational purposes, therefore, it was assumed that each subject considered each item, and that items which were left unanswered were not known by the subject. This assumption is somewhat risky in the case of the arithmetic examination since the last 20 items were the most complex on the examination and were answered only by a few subjects. It was, therefore, not possible, from a consideration of the tests alone, to determine whether an item had been reached by the subject and purposely left blank or if the item had not been attempted. The decision to compute item difficulties based on the assumption that each S considered each item was founded upon two observations: (1) Numerous Ss attempted items 1 through 20 (addition and subtraction), left items 21 through 40 blank (multiplication and division), attempted items 41 through 60 (addition and subtraction), and left items 61 through 80 (multiplication and division) blank. Since it is evident that Ss had time to consider items 21 through 40 but did not make written responses, it can be argued that written responses on more complex items of the same nature (multiplication and division) should not be expected. (2) The investigator circulated about the room during the examinations and observed subjects. He observed some Ss attempting items near the end of the exam and leaving the items blank.

In the case of the spelling examination where items were presented verbally, it can be assumed that each S had an equal opportunity (15 seconds) to attempt each item.

Selection of a Type of Task for the Main Study. Inspection of pilot test results suggested that spelling met all the specified requirements for a main study task. It was decided not to use arithmetic items because pilot study data revealed considerable variation in item difficulties according to the mathematical process involved. On every arithmetic item the public school students attained a higher proportion of correct responses than was attained by the institutionalized subjects. The latter group appeared particularly handicapped on multiplication and division items; nineteen of the 48 institutionalized subjects (39.8%) failed to answer any multiplication or division item correctly. Since the treatment procedure in the main study was not planned to incorporate an instructional process, the use of multiplication and division items seemed inappropriate, i.e., there was nothing within that testing session which would teach subjects to multiply or divide if they did not already know how. On the other hand, subjects were familiar with the process of placing letters together to form words. Although many of the words presented were not spelled correctly during the pilot testing phase, Ss did know the general process of connecting letters and, therefore, in the investigator's judgment, spelling seemed the more appropriate kind of task to employ in the main study phase.

Assignment of Subjects to Treatment Groups. Forty non-institutionalized subjects were randomly selected from among available pilot subjects and were randomly assigned to one of four treatment groups:

Group A: Easy Task--knowledge of past performance plus statement of "goal"
Group B: Hard Task--knowledge of past performance plus statement of "goal"
Group C: Easy Task--knowledge of past performance only
Group D: Hard Task--knowledge of past performance only

Forty institutionalized subjects were similarly randomly selected from among available pilot subjects and were randomly assigned to one of four treatment groups:

Group E: Easy Task--knowledge of past performance plus statement of "goal"
Group F: Hard Task--knowledge of past performance plus statement of "goal"
Group G: Easy Task--knowledge of past performance only
Group H: Hard Task--knowledge of past performance only

Pilot subjects who answered fewer than 5 items correctly on the pilot spelling test (i,e., those for whom an "easy" task could not be constructed) were excluded from the sample before random selection of main study subjects was made. Similarly subjects missing less than 10 items (i.e., those for whom a "hard" task could not be constructed) were eliminated from further testing.

Each treatment group contained ten (10) subjects.

Construction of Individual Tasks. On the basis of (1) item difficulties computed for pilot test data and (2) individual subject performance on the pilot test items, an individual task was constructed for each subject. A "hard" task was defined as a set of 10 items randomly selected from all the pilot test items between .33 and 0 difficulty level (DL) which were not answered correctly on the pilot test by the S for whom the items were being selected. In all cases DL was determined by the responses of pilot test subjects in the S's sub-sample, i.e., pilot test data from institutionalized subjects was used in constructing tasks for institutionalized main study subjects and pilot test data for non-institutionalized pilot test subjects was used in constructing tasks for non-institutionalized main study subjects.

Each S receiving a "hard" task, then, was given 10 items which he had been unable to answer correctly during pilot testing. The task was "hard", then, both in terms of the S's own demonstrated ability and in terms of the performance of the sub-sample to which S belonged.

An "easy" task was defined as a set of 10 items selected as follows for each subject: (a) 5 items randomly chosen from all the pilot test items between .33 and 0 DL which were not answered correctly on the pilot test by S, and (b) 5 items randomly selected from all the items answered correctly by S on the pilot test. "Easy" tasks differed from "hard" tasks, therefore, both in terms of the number of items known to S at the onset and in terms of the Mean Difficulty Level (MDL) (see Appendix D for a listing of the MDL of items selected for each S).

Standardization of Instructions and Experimental Procedure

Once items had been selected on the basis of group testing data and tasks had been constructed, it was still necessary to make sure that the instructions to be given in the individual testing sessions were clear and understandable to the subjects. It was also necessary for the experimenters to become familiar with the manipulation of the materials and the scoring procedure before testing any regular subjects. For these reasons, three pilot subjects were selected at random and given a complete series of tasks. Minor modifications in instructions were made on the basis of these "trial runs".

Although it was felt that three pilot subjects were an insufficient number to assure that instructional procedures had been sufficiently tested, pressing time demands led the investigator to begin testing main study subjects with the idea that, if the testing procedures proved to be inadequate, revisions would be made and new subjects selected to replace those already tested. Fortunately, no revisions were found to be necessary.

Main Study

Subjects

Subjects used in the main study were selected from those used in the pilot study. Appendix D presents a complete list of these subjects, including their MDL on spelling items selected, score on pilot tests, chronological age, mental age, IQ score and institutional-non-institutional status. All subjects were males. Means and standard deviations are also presented in Appendix D. Table 3.1 summarizes this data for non-institutionalized subjects, institutionalized subjects, and total for those subjects selected for the main study with regard to CA, MA, and IQ.

TABLE 3.1

Means and Standard Deviations on CA, MA, and IQ for Non-Institutionalized, Institutionalized, and Total Main Study Subjects

	CA		MA		IQ	
	Mean	SD	Mean	SD	Mean	SD
Non-Institutionalized N = 39*	188.89	17.19	135.97	30.80	80.72	14.09
Institutionalized N = 40	203.10	20.47	121.90	22.39	70.77	8.32
Total N = 79	194.67	20.68	128.85	27.76	75.68	12.56

*Information not available for one (1) non-institutionalized subject

Materials

Each test item word for the main study was typed in "primary" size black letters (¼ inch in height) on a 3 x 5 inch white card. Appendix E presents a sample of the stimulus cards used.

Record forms (see Appendix F) were especially prepared upon which the experimenter could list (1) the performance of each subject upon each trial, and (2) the goal set by each experimental subject before each new trial.

Experimenters

All testing sessions of the main study were conducted by one of three experimenters: (1) Dennis A. Warner, the Principal Investigator of the present research project and a doctoral student in Educational Psychology at the University of Oregon, (2) John C. Rolland, USOE Doctoral Fellow in Special Education at the University of Oregon and former Speech Therapist at Parsons State Hospital, Parsons, Kansas, and (3) Jack W. Martin, USOE Doctoral Fellow in Special Education at the University of Oregon and former clinical psychologist at Parsons State Hospital, Parsons, Kansas.

The use of three experimenters instead of only one has several obvious advantages. First, it facilitates the examination of the total sample of subjects in one-third the time. Second, it permits someone other than the investigator to get a close look at the experimental procedures and to detect possible weaknesses which may have been overlooked by him. Third, the use of three experimenters is a safeguard against the possible conscious or unconscious manipulation of factors by the investigator to produce the expected findings. There is ample research evidence to support the theory that emotionally involved experimenters unwittingly produce results they expect to occur. The use of two more "emotionally detached" experimenters should correct, in part, this tendency.

While there are some obvious advantages to the use of more than one experimenter, it should not be overlooked that there is at least one possible influence which may be viewed, by the individual researcher, as a disadvantage. Differences between experimenters may tend to increase the size of the error term in the analysis of variance formula. As Winer (1962) has explained:

> All uncontrolled sources of variance influencing an observation under a specified treatment combination contribute to what is known as the variance due to experimental error.... Differences between units of the experimental material existing prior to the experimental treatment, variance introduced by inaccuracies or uncontrolled changes in experimental techniques, possible unique interactions between the material and the treatments--all these sources contribute to the within-cell variance (p. 150).

It is very logical to believe that in investigations such as the present, where pre-established instructions are so few and where E's ability to relate to mentally retarded youth, to put them at ease, to communicate instructions, etc., is so important, differences between experimenters would be large enough to appreciably increase the error term in the analysis of variance. Since the error variance is used as the denominator in forming each F ratio in the particular design used in the present study, increases in this term would result in smaller F ratios.

White the influence of multiple experimenters upon the error variance may be viewed as a disadvantage to the individual investigator who is eager to confirm his hypotheses, this influence, in the long run, may eliminate many confirmations which, in fact, result from the idiosyncrasies of a single experimenter.

It was felt by the investigator that the advantages of using more than one experimenter outweighed the disadvantages and, therefore, Mr. Rolland and Mr. Martin were employed to assist in the data collection phase of the present investigation.

Testing Environment

Testing for the main study was conducted on an individual basis. In each case the testing of a single subject was conducted by one experimenter. Where facilities permitted each S to be tested alone in a small room this arrangement was used. In the case of Kennedy Junior High and the Special Education Center, it was necessary for all experimenters to work in the same classroom (Kennedy) or gymnasium (S.E.C.). In the latter case, distance between Ss was maximized and exchange between Ss was not permitted during testing. In most cases it was judged by Es that Ss could not hear the testing of other Ss.

During his test session, the subject sat at a desk and wrote his responses with a pencil upon yellow paper. The experimenter was seated to the left of right-handed subjects and to the right of left-handed subjects. This permitted his reading the response. Each item word was twice read aloud by E; readings were approximately 2 seconds apart. S was given 15 seconds after the second reading to write his response. At the end of the response, or at the end of 15 seconds, whichever came first, the card containing the correct spelling was placed in front of the S near his paper for 5 seconds. When the response was correct, E said "good" as he put down the card. When the response was not correct no verbal expression was made by E.

S's performance on each trial and, in the case of goal setting subjects, the estimates of future trials were recorded by E on a specially prepared form (see Appendix F).

Instructions to Subjects

After being seated at the desk, the subject was given the following instructions:

> "I have a list of 10 words that I would like you to spell. Each of the words is written on one of these cards (shows S pack of cards). I will read each word twice and then you will have 15 seconds to spell it on this paper (points) using the pencil (points). After you have written the word, I will show you the card so you can compare your spelling with the correct spelling to see if you got it right."
>
> "Let's try the first problem."

After S had attempted each item in the task, E said:

> "You got (number) out of 10 correct that time. Let's try it again."

At this point experimental Ss were also asked:

> "How many out of the same 10 words do you think you can get correct this time?"

These same instructions were given at the end of each of the first 9 trials. At the end of the 10th trial, control Ss were told how well they had performed and were thanked for participating. At the end of trial 10, experimental subjects were told:

> "You got (number) out of 10 correct that time. We are not going to spell the words again, but if we were, how many do you think you could get right next time?"

In the event that a S correctly responded to 10 out of 10 items on 3 consecutive trials, he was tested no further. In other cases, Ss received 10 trials each.

On occasion during the main study trials, a goal-setting subject would be reluctant to predict a performance level for the next trial. Such a S would often say, "I don't know how many I can get right" or something similar. In such a case E said: "How many do you think you can get right?" In no case was S allowed to continue until a prediction had been made. Some Ss had to be asked as many as three times to state a goal. This reluctance to express a goal generally occurred after the first trial (at the first time goal setting was requested) and seldom recurred once the S had made one prediction.

A somewhat related problem occurred when a control (non-goal setting) S stated a goal without being asked to do so. This occurred with only one S and, since it happened early in the testing session, he was asked by E not to verbally announce such goals even if they were made privately. This brings up a very interesting point related to this and similar studies of goal setting--Does the fact that control Ss do not express performance goals mean that they do not set personal goals during the testing sessions? It seems quite plausible that control Ss may look toward a level of performance which they consider appropriate and then work toward it without ever expressing such a goal to E. Perhaps the present study is really concerned with a difference between the performance of subjects with "publicly announced" and subjects with "privately determined" goals. Since no evidence was obtained regarding the prevalence of "private" goal setting, however, the present study will be described in terms of the original construct (i.e., goal setting and non goal setting) with the realization that some private goal setting may have occurred in the non goal setting groups.

Chapter 4

RESULTS

Testing of Major Hypothesis

English and English (1958) have clearly and concisely defined analysis of variance as:

> A method for determining whether the differences expressed as variance) found in a dependent variable, when it is exposed to the influence of one or more experimental variables, exceed what may be expected by chance. The F test is a measure of the probability of the beyond-chance difference (p. 28).

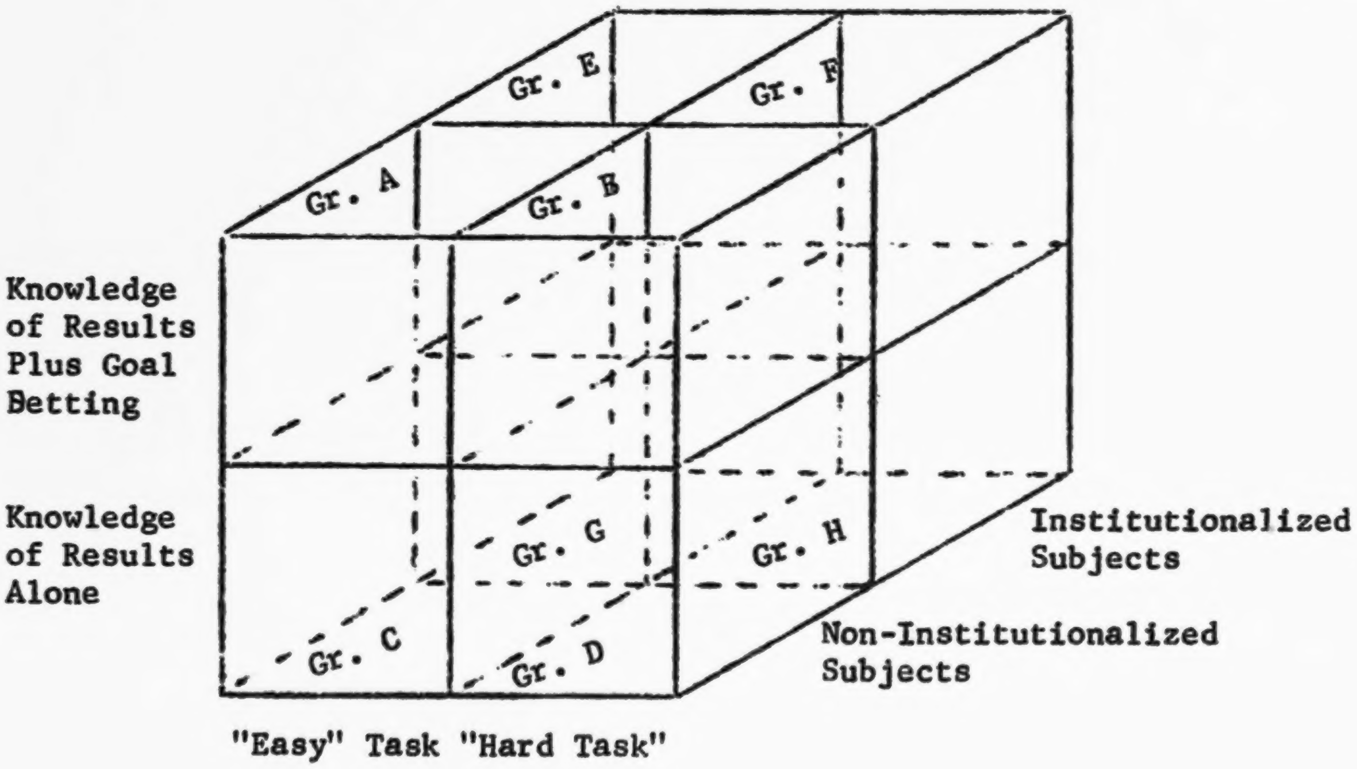
In the present study the effect of three experimental variables (goal setting, task difficulty level, and institutionalization) upon one dependent variable (task performance) were of primary concern. A 2 x 2 x 2 analysis of variance design, as depicted in figure 4.1, was, therefore, selected for use in the analysis of criterion (10th trial*) scores for the 80 main study subjects (see Appendix G for a complete list of scores for each S on each trial).

---

*In the event a subject attained three consectuive trials with a perfect score prior to Trial #10, his last score was used as his criterion score.

Figure 4.1

Schematic Representation of 2 x 2 x 2 Analysis of Variance Design Used in the Present Study


Gr. A
Gr. B
Gr. C
Gr. D
Gr. E
Gr. F
Gr. G
Gr. H
Knowledge of Results Plus Goal Setting
Knowledge of Results Alone
Institutionalized Subjects
Non-Institutionalized Subjects
"Easy" Task "Hard Task"

The design diagrammed in Figure 4.1 corresponds to the factor x factor x level (A x B x L) design proposed by Lindquist (1953, pp. 239-243) with the two factors being "goal setting vs. non goal setting" (factor A) and "'easy' task vs. 'hard' task" (factor B). "Institutionalization vs. non-institutionalization" constitutes the levels (L) of the design.

The computation of the A x B x L analysis of variance is presented in Appendix H. Table 4.1 shows the results of these computations.

TABLE 4.1

Summary Table of Analysis of Variance
Performed on Criterion (10th Trial) Scores

Sources of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F Ratio
Factor A	1	43.52	43.52	6.40*
Factor B	1	49.62	49.62	5.61*
Levels (L)	1	5.52	5.52	.71
AB	1	12.02	12.02	1.55
AL	1	.02	.02	0
BL	1	.02	.02	0
ABL	1	4.50	4.50	.58
w cells	72	557.70	7.75	
TOTAL	79	672.89		

*significant at .05 level
**significant at .01 level

Results of the A x B x L analysis of variance on criterion scores indicates a significant main effect for both factors A and B. To paraphrase the language of Edwards (1960, p. 188-189) we can say that the significant A mean square tells us that the means for A_1 (goal setting) and A_2 (non goal setting) averaged over the levels of B and L differ significantly. Examination of the means indicates that subjects who set goals attain a higher performance level, in general, than subjects who do not set goals.

Similarly, the significant B mean square tells us that the means for B_1 ("easy" task) and B_2 ("hard" task) averaged over the levels of A and L differ significantly. Examination of these two means indicates that subjects who perform "easy" tasks attain higher performance scores, in general, than do subjects who perform "hard" tasks. This finding, of course, is not surprising since subjects with "easy" tasks began their trials with approximately one-half their items already known (on the basis of pilot test results).

The fact that L mean square did not reach significance indicates that the means of L_1 (non-institutionalization) and L_2 (institutionalization) averaged over all levels of A and B do not differ significantly. Inspection of these means indicates that non-institutionalized subjects tend to attain higher performance scores than institutionalized subjects but that the difference is small enough to be accounted for by chance alone.

The A x B interaction is not significant, indicating that the higher scores obtained under the goal setting conditions maintains for both the "hard" and "easy" task conditions.

The A x L interaction is also not significant, indicating that the higher scores obtained under the goal setting condition maintains for both the institutionalized and non-institutionalized subjects.

The non significant B x L interaction simply indicates that both institutionalized and non-institutionalized subjects receiving the more difficult tasks earn lower scores than do subjects receiving the easier tasks.

The fact that the triple interaction, A x B x L, is not significant--paraphrasing Edwards (1960) once more--means that the A x B interactions for the separate levels of L are of the same form; that the A x L interactions for the separate levels of B are of the same form; and that the B x L interactions for the separate levels of A are of the same form.

Results of the 2 x 2 x 2 analysis of variance on criterion (10th trial) scores permits us to test our major hypothesis. This hypothesis stated:

> H_1: The mean performance score on the criterion measure (10th trial) for subjects who state goals will be greater than the corresponding mean performance score for subjects who only receive information on past performance.

The significant main effect for factor A permits us to retain H_1.

Testing of Auxiliary Hypotheses

Figure 4.2 schematically depicts the several treatment groups and reports their respective means.

As one would expect from the results of the 2 x 2 x 2 analysis of variance on criterion scores, the group means represented in Figure 4.2 indicate a general tendency for goal setting groups to exceed non goal setting groups and for groups performing "easy" tasks to attain a higher score than that of groups performing "hard" tasks. The significance of differences between pairs of treatment means must, however, be determined through appropriate statistical procedures.

Figure 4.2

Arithmatic Mean of Each of Eight Treatment Groups on Criterion (10th Trial) Scores. (N = 10 for Each Group)

Non-Institutionalized Subjects

	"Easy" Task	"Hard" Task
Knowledge of Results plus Goal Setting	A 4.7 8.1 3.4	B 1.9 7.8 5.8
Knowledge of Results Alone	C 4.3 7.9 3.6	D .6 5.1 4.5

Institutionalized Subjects

	"Easy" Task	"Hard" Task
Knowledge of Results plus Goal Setting	E 4.1 8.1 4.0	F 1.0 6.8 5.8
Knowledge of Results Alone	G 4.7 6.9 2.2	H 1.0 5.0 4.0

Regarding the testing of differences between means, common practice in past years has been, once an F was found significant, to do a series of t tests or to use the least significant difference method (LSD-method). These methods, in general, are felt to give a high degree of false significances and are, therefore, gradually being abandoned (Seeger, 1966).

Edwards (1960) has compared a number of multiple comparison techniques. Of these methods he has written:

> In making multiple comparisons among the treatment means, it is not necessary that the treatment mean square of the analysis of variance be significant. In other words, we may have a nonsignificant treatment mean square and still use the methods...for making multiple comparisons. This does not mean, however, that we

> should indiscriminately apply the methods, one after another, with the anticipation that one or another may result in some finding that meets the requirements of statistical significance. Our choice of methods should, instead, be guided by questions of experimental interest (p. 136).

For the present study the investigator elected to use Duncan's New Multiple Range Test (Duncan, 1955) to test the significance of the difference between pairs of means.

TABLE 4.2

Duncan's New Multiple Range Test Applied to Criterion Score Means for 8 Treatment Groups

Means	(1) H 5.0	(2) D 5.1	(3) F 6.8	(4) G 6.9	(5) B 7.8	(6) C 7.9	(7) E 8.1	(8) A 8.1	(9) Shortest Significant Ranges*
H 5.0		.1	1.8	1.9	2.8	2.9	3.1	3.1	R_2 = 2.48
D 5.1			1.7	1.8	2.7	2.8	3.0	3.0	R_3 = 2.61
F 6.8				.1	1.0	1.1	1.3	1.3	R_4 = 2.70
G 6.9					.9	1.0	1.2	1.2	R_5 = 2.76
B 7.8						.1	.3	.3	R_6 = 2.81
C 7.9							.2	.2	R_7 = 2.85
E 8.1								0	R_8 = 2.88
A 8.1									

H D F G B C E A

*at .05 level

Any two treatment means not underscored by the same line are significantly different.

Any two treatment means underscored by the same line are not significantly different.

Results of the Duncan's New Multiple Range Test permit us to test each of the auxiliary hypotheses. These hypotheses were stated as follows:

> H_2: Non-institutionalized, "Hard" Task, Goal Setting subjects will have a greater mean performance score on

the criterion measure (10th trial) than the corresponding mean performance score of Non-institutionalized, "Hard" Task, Non-Goal Setting subjects.

H_3: Non-institutionalized, "Easy" Task, Goal Setting subjects and Non-institutionalized, "Easy" Task, Non-Goal Setting subjects will not differ with respect to mean performance on the criterion measure (10th trial).

H_4: Institutionalized, "Hard" Task, Goal Setting subjects will have a greater mean performance score on the criterion measure (10th trial) than the corresponding mean performance score of Institutionalized, "Hard" Task, Non-Goal Setting subjects.

H_5: Institutionalized, "Easy" Task, Goal Setting subjects and Institutionalized, "Easy" Task, Non-Goal Setting subjects will not differ with respect to mean performance on the criterion measure (10th trial).

The significant difference between the mean of group B and the mean of group D permits us to retain H_2.

The non-significant difference between the mean of group A and the mean of group C permits us to retain H_3.

The non-significant difference between the mean of group F and the mean of group H forces us to reject H_4.

The non-significant difference between the mean of group E and the mean of group G permits us to retain H_5.

Discussion of Findings

The most important finding of the present research is, undoubtedly the fact that the mean performance of subjects who set goals was significantly larger than the mean performance of subjects who did not set goals. This finding is in harmony with those of Armstrong (1947), Fryer (1964), Lockette (1956), and Kausler (1959) (see Chapter 2) and suggests that goal setting, for mentally retarded male adolescents, has a motivational property beyond that of mere feedback information alone. It must be kept in mind, however, that the present research project dealt with a very specific kind of task (i.e., spelling) and that the conditions under which the task was presented were, likewise, of a specific nature. It would be dangerous to assume, therefore, that mentally retarded male adolescents would perform similarly under different experimental conditions and/or upon a different kind of task. The findingsof the present project suggest a real need for further research employing different tasks and various testing conditions. Since goal setting may have implications for classroom teaching, future research should investigate the effects of goal setting upon actual classroom performance. Such research, for example could involve a number of randomly assigned MR classes where the teacher asked each student to estimate his daily (or weekly) progress on a number of class assignments (arithmetic problems, list of spelling words, new vocabulary, etc.) and then compare with him his actual progress at the end of the pre-established period of time. Estimates and actual performance could even be recorded upon a specially prepared chart for general display.

A similar number of classes where no special goal setting techniques were arranged would serve as a control group. Differences between pre-tests and post-tests could serve as a measure of progress for each child. Such a procedure would yield information regarding the effects of goal setting upon classroom performance. Such a project would certainly be an ambitious one and, undoubtedly, a costly one but the answers obtained could have significant and profound implications for education.

The findings related to auxiliary hypotheses 2 and 3 suggest that goal setting among public school EMR males has a pronounced effect upon subsequent performance on a "hard" task but not on an "easy" one. Within the public school sub-sample the mean performance score of the "Hard" Task, Goal Setting subjects was found to be significantly greater ($P = .05$) than the mean performance score of the "Hard" Task, Non-Goal Setting group. Within the same sub-sample, however, there was not a significant difference between the mean performance scores of Goal Setting and Non-Goal Setting subjects when the "easy" task was considered alone. These results are in harmony with those of Fryer (1964) as reviewed in Chapter 2.

Within the institutionalized sub-sample there was a tendency for "Hard" Task, Goal Setting subjects to attain a greater criterion score than that attained by "Hard" Task, Non-Goal Setting subjects. This difference, although in the predicted direction, did not, however, reach significance at the .05 level. Likewise, as predicted, the mean score of the "Easy" Task, Goal Setting subjects did not differ significantly from the mean score of the "Easy" Task, Non-Goal Setting subjects.

In summary, it can be said that goal setting has a greater effect upon public school EMRs than upon institutionalized EMRs as judged by performance on tasks of a difficult nature. This finding may suggest that the use of goal setting in the regular classroom instructional process may be more effective in public school EMR classes than in institutional EMR classes. More research seems warranted, however, before such a conclusion is accepted for actual application.

A finding, unrelated to the state hypotheses but which merits some consideration, is that within the non-institutionalized sub-sample there was a significant difference between the mean performance score of the "Easy" Task, Non-Goal Setting subjects and the mean performance score of the "Hard" Task, Non-goal Setting subjects. However, the mean performance scores for "Easy" Task subjects and "Hard" Task subjects did not differ significantly among those who set goals. In other words, it could be argued that goal setting behavior increases performance on the "Hard" task to a greater extent than it increases performance on the "easy" task. This increase brings the mean of the "Hard" Task, Goal Setting subjects very close to the mean of the "Easy" Task, Goal Setting subjects in numerical value. Among the non-goal setting subjects, however, the mean of the "Hard" task subjects remains significantly smaller than the mean of the "Easy" task subjects.

Within the institutionalized sub-sample there was a tendency for means to show the above discussed relationships. However, differences between the mean performance score of the "Easy" Task, Non-Goal Setting subjects and the mean performance score of the "Hard" Task, Non-Goal Setting subjects, although in the predicted direction, did not attain significance at the .05 level.

Chapter 5

SUMMARY AND IMPLICATIONS

Problem

Several investigators, in studying non-retarded subjects, have concluded that goal setting has a motivational property beyond that of mere feedback information, i.e., task performance scores are greater under conditions where subjects state a "level of aspiration" before each trial than under conditions where subjects do not. The purpose of the present study was to investigate the effects of goal setting upon the performance of educable mentally retarded male adolescents on educational tasks of differing degrees of difficulty. A comparison was also made between institutionalized and non-institutionalized subjects with regard to the effects of goal setting and task difficulty upon performance.

More specifically, the following major hypothesis was tested in the research project:

> H_1: The mean performance score on the criterion measure (10th trial) for subjects who state goals will be greater than the corresponding mean performance score for subjects who only receive information on past performance.

In addition to the major hypothesis several auxiliary hypotheses were also tested:

> H_2: Non-institutionalized, "Hard" Task, Goal Setting subjects will have a greater mean performance score on the criterion measure (10th trial) than the corresponding mean performance score of Non-institutionalized, "Hard" Task, Non-Goal Setting subjects.

> H_3: Non-institutionalized, "Easy" Task, Goal Setting subjects and Non-institutionalized, "Easy" Task, Non-Goal Setting subjects will not differ with respect to mean performance on the criterion measure (10th trial).

> H_4: Institutionalized, "Hard" Task, Goal Setting subjects will have a greater mean performance score on the criterion measure (10th trial) than the corresponding mean performance score of Institutionalized, "Hard" Task, Non-Goal Setting subjects.
>
> H_5: Institutionalized, "Easy" Task, Goal Setting subjects and Institutionalized, "Easy" Task, Non-Goal Setting subjects will not differ with respect to mean performance on the criterion measure (10th trial).

These auxiliary hypotheses were suggested by the finds of Fryer (1964) which indicate that goal setting is associated with greater performance scores than non-goal setting when a "difficult" task is involved but not when the task is "easy" for subjects to perform. Both institutionalized and non-institutionalized (public school) educable mentally retarded subjects were used in the present study so that finding could be generalized to both sub-populations.

Procedures

Pilot Study

The purpose of the pilot study was twofold: (1) the development of appropriate educational tasks for use in later testing, and (2) the standardization of instructions and experimental procedures.

Subjects. Ss for the pilot study were 93 educable mentally retarded male adolescents (CA range = 238 to 150 months). Forty-eight (48) of the subjects were from Fairview State Hospital and Training Center in Salem, Oregon; forty-five (45) of the subjects were from public school classes for the educable mentally retarded in the Eugene, Oregon and Salem, Oregon school districts.

Each pilot subject completed a group administered paper-and-pencil test containing 80 arithmetic and 80 spelling items. Items for the test were selected from a variety of test books and curriculum guides for mentally retarded junior high school and high school age students. The pilot study data was used to compute an item difficulty level for each of the 160 test items.

Individual Tasks. An individual task was constructed for each main study subject on the basis of (1) item difficulties computed from pilot test data and (2) individual S performance on the pilot test items. A "hard" task was defined as 10 items, each with a difficulty level (DL) between .33 and 0 and which were missed by the subject during group pilot testing. The 10 items were selected at random from all those between .33 and 0 difficulty which were missed by the subject for whom the task was being constructed.

An "easy" task was defined as one consisting of (a) 5 items selected at random from all those between .33 and 0 DL missed by the S on the pilot test and (b) 5 items selected at random from those answered correctly by S on the pilot test. In all cases a S's items were selected with regard to the pilot test performance of his own subsample.

The decision to define "easy" and "hard" tasks in the above manner was based upon the following considerations: (1) Each task should be one in which individual subject performance can increase as a function of repeated contact with the stimulus. (2) The type of task used should be of such a nature that the S knows the general manner in which the task is performed, i.e., the process which he must go through to successfully complete an item. (3) A "hard" task should differ from and "easy" task in terms of the subject's ability to perform at the onset and in terms of the complexity of the task itself. (4) Each individual's task, whether "hard" or "easy", should be approximately equal in difficulty to those of other persons within his respective treatment group.

On the basis of these specifications, the investigator decided to use spelling tasks in the main study. The reasons for settling upon spelling were as follows: (1) Such a task is one in which performance can increase as a result of practice, i.e., repeated experience with the stimulus word. (2) Ss were, as a general rule, familiar with the process of connecting alphabetic letters in a chain to form words. Arithmetic was abandoned as a possible basis for main testing because pilot test items tended to differ in difficulty according to the mathematical process involved, i.e., most addition and subtraction problems were between 1.0 and .33 DL and all multiplication and division items were below .33 DL. Since a number of subjects, primarily from the institution, answered no multiplication and division problems correctly, such tasks seemed inappropriate. There was nothing in the treatment designed to teach a process; the experimental procedure assumed that Ss already knew the process required to successfully complete an item and that repeated exposure to the stimulus would facilitate increased performance.

Each test item word for the main study was typed in "primary" size black letters (¼ inch in height) on a 3 x 5 inch white card.

Main Study

Subjects. Ss for the main study were 40 institutionalized ($\overline{CA}$ = 203.1 months; MA = 121.9 months; $\overline{IQ}$ = 70.8) and 40 special education students (CA = 186.0 months; $\overline{MA}$ = 136.0 months; $\overline{IQ}$ = 80.7) who were randomly selected from among the pilot study subjects.

Non-institutionalized subjects were randomly assigned to one of four treatment groups:

Group A:	Easy Task--knowledge of past performance plus statement of "goal"
Group B:	Hard Task--knowledge of past performance plus statement of "goal"
Group C:	Easy Task--knowledge of past performance only
Group D:	Hard Task--knowledge of past performance only

Institutionalized subjects were similarly randomly assigned to one of four treatment groups:

Group E:	Easy Task--knowledge of past performance plus statement of "goal"
Group F:	Hard Task--knowledge of past performance plus statement of "goal"
Group G:	Easy Task--knowledge of past performance only
Group H:	Hard Task--knowledge of past performance only

Each treatment group contained ten (10) subjects. In all a total of eighty (80) subjects were tested in the main study.

Testing. The 80 subjects in the main study testing group were tested alone by a single experimenter. The testers were three advanced graduate students of experience in testing retarded children. Ss were either tested alone in a small room or in discrete areas within a large assembly room in which the distance between Ss was maximized and verbal exchange between Ss prohibited during testing.

During his test session, each subject sat at a desk and wrote his responses with a pencil upon yellow paper. The experimenter was seated to the left of right-handed subjects and to the right of left-handed subjects. This permitted his reading the subject's written response. Each item word was twice read aloud by E; readings were approximately 2 seconds apart. S was given 15 seconds after the second reading to write his response. At the end of the response, or at the end of 15 seconds, whichever came first, the card containing the correct spelling was placed in front of the S near his paper for 5 seconds. When the response was correct, E said "good" as he put down the card. When the response was not correct, no verbal expression was made by E.

Instructions to Subjects. After being seated at the desk, the subject was given the following instructions:

> "I have a list of 10 words that I would like you to spell. Each of the words is written on one of these cards (shows S pack of cards). I will read each word twice and then you will have 15 seconds to spell it on this paper (points) using the pencil (points). After you have written the word, I will show you the card so you can compare your spelling with the correct spelling to see if you got it right."
>
> "Let's try the first problem."

After S had attempted each item in the task E said:

> "You got (number) out of 10 correct that time. Let's try it again."

At this point experimental Ss were also asked:

> "How many out of the same 10 words do you think you can get correct this time?"

These same instructions were given at the end of each of the first 9 trials. At the end of the 10th trial, control Ss were told how well they had performed and were thanked for participating. At the end of trial 10, experimental subjects were told:

> "You got (number) out of 10 correct that time. We are not going to spell the words again, but if we were, how many do you think you could get right next time?"

In the event that a S correctly responded to 10 out of 10 items on 3 consecutive trials, he was tested no further. In other cases, Ss received 10 trials each. The S's number of correct items on his 10th trial was used as his criterion score for the data analysis. The criterion score for Ss receiving less than 10 trials was the maximum score of 10.

The criterion (10th trial) scores (i.e., number of correct responses) were analysed in a three way (factor x factor x level) analysis of variance design with "goal setting" and "task difficulty" as the two factors and "institutionalization--non-institutionalization" constituting the levels. The main effects for both factors were found to be significant beyond the .05 level but the main effect of the levels was not significant at the .05 level. None of the interaction effects were significant at the .05 level.

Examination of the column means indicates that goal-setting is superior to non-goal-setting, i.e., higher scores are associated with goal-setting than with feedback information alone. Examination of the row means indicates that higher scores are associated with the "easy" task than with the "hard" task.

Differences between the mean scores of institutionalized and non-institutionalized subjects are accountable by chance variations.

The findings of this study regarding the effects of goal setting are generally in harmony with earlier studies employing non-mentally retarded subjects, i.e., that subjects who set performance goals attain a higher score, in general, than similar subjects who do not.

The finding that "easy" tasks are associated with higher scores than are "hard" tasks is not surprising. The concern of the present study was not so much with the main effects of task difficulty as with the relationship between difficulty and goal setting within each of the sub-samples. Likewise, since tasks for institutionalized and non-institutionalized subjects were constructed on the basis of their respective sub-sample performance on the pilot test, differences within sub-samples and not between them were of primary concern.

The four (4) auxiliary hypotheses (H_2 through H_5) examined using Duncan's New Multiple Range Test revealed that:

(1) the mean performance score of the Non-Institutionalized "Hard" Task, Goal Setting subjects was significantly greater (P = .05) than the mean performance score of the Non-Institutionalized "Hard" Task, Non-Goal Setting subjects.

(2) the mean performance score of the Non-Institutionalized, "Easy" Task, Goal Setting subjects was not significantly different (P = .05) than the mean performance score of the Non-Institutionalized, "Easy" Task, Non-Goal Setting subjects.

(3) that the mean performance score of the Institutionalized, "Hard" Task, Goal Setting subjects was not significantly greater (P = .05) than the mean performance score of the Institutionalized "Hard" Task, Non-Goal Setting subjects; and

(4) that the mean performance score of the Institutionalized, "Easy" Task, Goal Setting subjects was not significantly different (P = .05) than the mean performance score of the Institutionalized, "Easy" Task, Non-Goal Setting subjects.

In short, the expectation that goal setting would be superior to non-goal setting in terms of performance on "hard" task but not on an "easy" task was upheld for the public school sub-sample but not for the institutionalized sub-sample. Although the difference between the mean performance score of the Institutionalized, "Hard" Task, Goal Setting subjects and the mean performance score of the Institutionalized, "Hard" Task, Non-Goal Setting subjects was in the predicted direction, it did not reach significance at the .05 level.

Limitations of the Study

The fact that a number of subjects, particularly those perforing "easy" tasks, attained a perfect score (10 correct) prior to the final trial and could not, therefore, increase their score because of the low ceiling, limits conclusions with respect to the difficulty factor. More extensive research with respect to the relationship of task difficulty and goal setting is needed. One experimental approach here might be to involve subjects in a series of subtasks of varying difficulty levels, having them set separate goals for themselves on each of these subtasks and noting whether their task performance differs for differing difficulty levels as contrasted with control subjects making no estimates. Performance might be measured in terms of task persistance (time) or work rate to avoid scoring problems tied to perfect score ceilings.

A second limitation of the present study is that only one kind of educational task--spelling as learned by an individualized flash card procedure,was investigated. Generalization of findings to other subject matter areas can only be hazarded. While it may seem reasonable that goal setting would well enhance learning in other content areas such as arithmetic, word recognition, time telling, etc. the particular mode of presentation of stimulus and feedback could well be a crucial factor enhancing (or offsetting) the goal setting effect. Again data is needed from controlled experiments.

A further limitation of the study resides in the difficulty in securring "comparable" samples of institutionalized and non-institutionalized retardates. Even limiting himself to "matching" only on those more readily obtainable descriptions such as sex, age, and measured intelligence, availability of subjects restricts experimenter [illegible] freedom. A more crucial concern is the relevance of "available descriptions" as contrasted to the more difficult-to-obtain descriptions of prior educational treatment, teacher-pupil interaction, etc., etc. The fact that investigators match on what is "handy" in no way assures adequate matching. In the present study the public school samples had generally higher IQ's based on the Peabody Picture Vocabulary Test (an average of nearly ten IQ points favoring the public school sample) and were approximately fourteen months older on the average both chronologically and in terms of mental age. The importance of these differences as related to obtained differences (and similarities) between the institutionalized and non-institutionalized retardates in this study must be left to reader judgement.

Implications for Classroom Teaching

The value of any study resides in its implications for action be it reformulation of theory or recommendation for behavioral change. Although results of the present study suggest that male educable mentally retarded subjects have a tendency to attain a higher final performance under goal setting conditions than under conditions where they are not asked to verbally express an estimate of future performance, the fact that the data derives from a very specific kind of task and employed a specific type of testing procedure limits direct implication for classrom teaching. The classroom teacher, will seldom have the time to work with a single student for an extended block of time (45 to 60 minutes). Nor will spelling materials ever occupy a major portion of her curriculum concerns.

The implications for application of the present findings to the classroom setting are the recognition of some very interesting and exciting possibilities by the teacher who may want to "experiment" with goal setting among his own students. The study has revealed some positive effects of "goal setting" behavior obtained under experimenter controlled conditions. The demonstration needs extension to the daily, complex situation of the actual classroom. A multitude of possible, small, inobtrusive classroom "experiments" are ready for the doing, minor introductions of "goal setting" behaviors to facilitate learning. Having students estimate the number of arithmetic problems they can do in a given period of time, how long they can sit without disrupting the class, how many pages of reading they can finish in a given period of time, how many days of the week they can remember to bring paper and pencil to class, how many words from a given list they can spell, etc. etc.; these are only a few of the ways in which "goal setting" may be introduced within classes for the educable mentally retarded. The incorporation of such goal setting techniques into the regular educational process would appear to require only a small amount of additional time and effort on the teacher's part. Accurate record keeping would help her impartially determine for herself and in her own classroom setting whether or not these applications of "goal setting" bring about improved performance in her students. Improvization and extension appear to be almost without limit.

Implications for Future Research

The present research project is offered as a pioneering effort in the investigation of goal setting as an independent variable in the learning of retardates. Its results must be accepted with reservation and we look for future studies to corroborate or contradict our present conclusions. Not only is there a need for more research of the same nature as that described here (i.e., research using carefully selected tasks and concentrating upon individual verbal goal setting), but research dealing with actual classroom tasks in group situations should likewise be undertaken. The informal "experimentation" by classroom teachers suggested in the preceeding paragraph could be expanded to include a large number of classrooms in which the well described procedures are followed and in which careful pre- and post tests on performance are made.

A major problem encountered in the present study centered around the definition and construction of the learning tasks. Though hindsight sometimes favors the resolution of problems which were unforeseen by the investigator at the outset of the project, the problem of defining logically sound, non-artificial learning tasks yielding interpretable performance measures for goal setting-learning studies appears to have no generalizable solution. The precept that the task must appropriate for the experimention needs simply passes the buck to the experimenter. The description of problems encountered in the present research project with regard to the definition and construction of tasks perhaps will serve as a starting gate for that more definitive next study.

The possibility that the learning of educable mentally retarded students may be facilitated by goal setting has been strongly suggested by the results reported here. It is hoped that the present study will serve as the basis for extensive and fruitful research related to the effects of goal setting upon learning in mentally retarded children, adolescents, and adults.

BIBLIOGRAPHY

Andreas, B. G. Experimental psychology. New York: John Wiley, 1960.

Armstrong, D. E. Performance as a function of expressed and non-expressed levels of aspiration. Unpublished master's thesis, Howard University, 1947.

Baytan, J. S. Interrelations between levels of aspiration, performance, and estimates of past performance. J. Exp. Psychol., 1943, 33, 1-21.

Beloit (Wisconsin) Public Schools, Language Arts Curriculum Bulletin. Beloit, Wisconsin: Department of Exc. Educ., 1965.

Conrad, H. S. The experimental tryout of test materials. In E. F. Lindquist (ed.), Educational measurement, Washington, D.C.: Amer. Council on Educ., 1951, pp. 250-265.

Cromwell, R. L. A social learning approach to mental retardation. In N. R. Ellis (ed.), Handbook of mental deficiency. New York: McGraw-Hill, 1963, 41-91.

Dembo, Tamara. Der Anger als dynamishches Problem. Psychol. Forsch., 1931, 15, 1-144.

Department of Education, State of Alaska, Arithmetic Guide, Juneau, Alaska, 1967.

Drew, C. J., Prehm, H. J., & Logan, D. R. Studies in paired associate learning: II. Performance of retarded and normal subjects as a function of association value of materials. Regional Rehabilitation Research and Training Center in Mental Retardation Working Paper No. 5, University of Oregon, 1967.

Duncan, D. B. Multiple range and multiple F tests. Biometrics, 1955, 11, 1-42.

Edwards, A. L. Experimental design in psychological research, rev. ed. New York: Holt, Rinehart, and Winston, 1960.

English, H. B. & English, A. C. A comprehensive dictionary of psychological and psychoanalytical terms. New York: David McKay, 1958.

Eyeman, R. K. Covariation of level of aspiration and adaptation level with other characteristics. Amer. J. Ment. Defic., 1964, 68, 741-749.

Filter, R. O. Estimates of the amount of work one can do. J. Appl. Psychol., 1927, 11, 58-67.

Frank, J. D. Recent studies of the level of aspiration. Psychol. Bull., 1941, 38, 218-225.

Fryer, F. W. An evaluation of level of aspiration as a training procedure. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Gardner, J. W. The use of the term "level of aspiration." Psychol. Rev., 1940, 47, 59-68.

Glim, T. E. & Manchester, F. S. Teaching instructions for Basic Spelling Keys 2, 3, 4, 5, 6, & 7 (Revised Editions). Philadelphia: J. B. Lippincott, 1967.

Horrocks, Edna M., Evans, Edith M., & Staiger, Ralph C. Spelling: Grade Two; Grade Four; Grade Five; and Grade Six. Boston: Ginn and Co., 1965.

Hutt, M. L. & Gibby, R. G. The mentally retarded child: Development, education, and guidance. Boston: Allyn and Bacon, 1958.

Jastak, J. & Bijou, S. Wide range achievement test. 1946 edition.

Kausler, D. H. Aspiration level as a determinant of performance. J. Pers., 1959, 27, 356-361.

Kirk, S. A. Educating exceptional children. Boston: Houghton Mifflin, 1962.

Kneeland, N. Self-estimates of improvement in repeated tasks. Arch. Psychol., 1934, 163, 1-75.

Lindquist, E. F. Design and analysis of experiments in psychology and education. Boston: Houghton Mifflin, 1953.

Lockette, R. E. The effects of level of aspiration upon the learning of skills. Dissertation Abst., 1956, 16, 284 (Abstract).

May, M. & Doab, L. Competition and cooperation. New York: Social Science Research Council, Bull. 25, 1947.

Morgan, C. T. Introduction to psychology. New York: McGraw Hill, 1956.

Purdom, G. A., Jr. Certification of educable mentally retarded children for special class placement in Oregon. Paper presented at the 87th National Meeting, AAMD, Portland, Oregon, May 22, 1963.

Robinson, H. B. & Robinson, N. M. The mentally retarded child. New York: McGraw Hill, 1965.

Seeger, P. Variance analysis of complete designs: Some practical aspects. Stockholm, Sweden: Almqvist & Wiksell, 1966.

Subotnik, L. Level of aspiration, emotional disturbance, and learning in institutionalized educable mentally retarded boys. *Am. J. Ment. Deficiency*, 1967, 71, 767-771.

The Board of Public Education, Pittsburgh, Pennsylvania, *Course of study for special education*, Grade 7, Grade 8, Grade 10. Pittsburgh, Pa.: Nowember, 1963.

Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

Yacorzynski, G. K. Degree of effort. III. Relationship to the level of aspiration. *J. Exp. Psychology*, 1942, 30, 407-413.

APPENDIX A

List of School, CA, MA, and IQ, for Non-Institutionalized and Institutionalized Pilot Test Subjects

Non-Institutionalized Subjects

Subject Number	School	CA*	MA**	IQ**
001	JEK	171	195	112
002	JFK	150	185	111
003	JFK	169	183	106
004	JFK	192	206	104
005	JFK	167	171	101
006	JFK	196	187	98
007	JFK	172	167	96
008	JFK	195	183	96
009	JFK	172	151	93
010	JFK	158	141	91
011	JFK	188	155	87
012	JFK	197	151	85
013	JFK	187	141	84
014	SEC	183	132	83
015	SEC	188	132	83
016	JFK	187	141	81
017	SEC	189	141	81
018	NSH	186	125	79
019	JFK	179	125	79
020	JFK	186	132	79
021	NSH	212	149	78
022	JFK	164	120	77
023	SEC	178	122	77
024	JFK	174	120	76
025	NSH	197	145	76
026	NSH	212	145	76
027	SEC	175	113	75
028	JFK	166	110	73
029	SEC	176	113	73
030	SEC	197	122	72
031	SEC	188	122	72
032	SEC	168	107	72
033	NSH	220	130	72
034	NSH	222	130	72
035	JFK	189	116	70
036	NSH	220	127	70
037	NSH	200	127	70
038	NSH	221	127	70
039	SEC	182	101	68
040	SEC	196	110	68

Appendix A Continued

Subject Number	School	CA*	MA**	IQ**
041	SEC	173	87	61
042	NSH	205	99	60
043	NSH	193	85	57
044	NSH	196	78	56
045	JFK	---	--	--
Mean		187.18	135.20	80.00
Standard Deviation		17.04	29.21	13.47

Institutionalized Subjects

Subject Number	CA*	MA**	IQ**
101	157	145	86
102	238	167	84
103	193	149	84
104	183	132	83
105	234	160	82
106	203	157	82
107	183	172	81
108	181	125	79
109	187	125	79
110	216	145	78
111	198	128	77
112	205	136	77
113	222	145	76
114	220	144	76
115	217	120	76
116	195	116	74
117	211	130	74
118	185	113	73
119	226	132	73
120	221	132	73
121	209	128	73
122	201	120	71
123	177	105	71
124	198	120	71
125	178	105	70
126	162	99	70
127	220	125	79
128	222	129	68

Appendix A Continued

Subject Number	CA*	MA**	IQ**
129	219	124	68
130	177	105	68
131	181	97	66
132	195	105	66
133	231	148	65
134	221	120	65
135	202	110	65
136	206	103	62
137	223	105	61
138	218	105	61
139	174	87	61
140	179	87	61
141	185	112	60
142	153	85	70
143	202	94	57
144	199	85	57
145	230	97	57
146	230	96	56
147	229	92	55
148	167	79	53
Mean	201.32	119.25	69.88
Standard Deviation	21.78	22.93	8.62

*CA at time of present study.

**MAs and IQs obtained from Peabody Picture Vocabulary Test Form A. Special class placement was based on various individually administered intelligence tests administered several years previously. The higher Peabody IQs obtained by several of the subjects suggest both a "test leniency" and/or actual increased "intelligence" during the retest interim.

APPENDIX B

Copy of Booklet Used in Pilot Testing

NAME ____________________

SCHOOL ____________________

ARITHMETIC

ROW 1* ADD	7 +2	9 +5	14 +7	1 +1	5 +9	6 +2	13 +7	1 +3	4 +5	6 +3
ROW 2 SUBTRACT	9 -2	10 -9	12 -7	4 -1	17 -12	9 -7	14 -12	5 -3	10 -1	40 -10
ROW 3 MULTIPLY	2 X 1	4 X 2	7 X 7	21 X 2	10 X 2	15 X 2	12 XX 3	6 X 2	7 X 4	9 X 9
ROW 4 DIVIDE	2)4	7)21	3)12	2)6	4)8	8)64	9)81			
	5)25	3)18								
ROW 5 ADD	10 +9	17 +12	21 +24	13 +13	109 +27	123 +49	270 +11	100 +100	92 +14	121 +21

*Items in Row 1 are referred to as items 1, 2, 3...10 in order of presentation; items in Row 2 are referred to as items 11, 12, 13... 20; etc.

ROW 6	14	21	23	10	20	47	92	102	7L	187
SUBTRACT	- 3	-20	-11	-3	-6	-13	- 7	-43	-12	-153

ROW 7	17	10	14	12	120	25	40	17	181	72
MULTIPLY	X 3	X 9	X 11	X10	X 4	X 3	X 4	X 10	X 2	X 9

ROW 8
DIVIDE $9\overline{)72}$ $20\overline{)40}$ $72\overline{)154}$ $7\overline{)53}$ $12\overline{)120}$ $19\overline{)57}$

$8\overline{)32}$ $12\overline{)144}$ $6\overline{)24}$ $5\overline{)1000}$

SPELLING

1. (A)**
2. (B)
3. (O)
4. (S)
5. (E)
6. (R)
7. (T)
8. (H)
9. (in)
10. (see)
11. (cat)
12. (milk)
13. (tree)
14. (eat)
15. (ran)
16. (man)
17. (fun)
18. (sit)
19. (hot)
20. (little)
21. (book)
22. (boy)
23. (box)
24. (thank)
25. (room)
26. (toy)
27. (birthday)
28. (leg)
29. (must)
30. (money)
31. (moon)
32. (Indian)
33. (ear)
34. (dish)
35. (party)
36. (nest)
37. (breakfast)
38. (lessons)
39. (monkey)
40. (snowball)
41. (turkey)
42. (Saturday)
43. (forest)
44. (lion)
45. (pocket)
46. (sudden)
47. (judge)
48. (nurse)

**Spelling words have been listed on the sample booklet, but, of course, did not appear on the booklets actually used in pilot testing.

49. (peanuts)
50. (piano)
51. (circle)
52. (history)
53. (wolf)
54. (apron)
55. (nickel)
56. (motor)
57. (polite)
58. (tiny)
59. (speed)
60. (ribbon)
61. (doctor)
62. (spider)
63. (orange)
64. (truth)
65. (animal)
66. (electric)
67. (hero)
68. (idea)
69. (spectator)
70. (deposit)
71. (convert)
72. (brother)
73. (sister)
74. (glass)
75. (iron)
76. (wrong)
77. (eaten)
78. (market)
79. (study)
80. (cereal)

APPENDIX C

List of Item Difficulties for Institutionalized and Non-Institutionalized Pilot Subjects on Each Spelling and Aritmetic Item

Spelling			Arithmetic		
Item No.	Difficulty N-Inst.	Difficulty Inst.	Item No.	Difficulty N-Inst.	Difficulty Inst.
1	1.00	.96	1	.91	.90
2	1.00	.90	2	.95	.79
3	1.00	.88	3	.84	.65
4	1.00	.90	4	1.00	.92
5	.99	.85	5	.95	.77
6	.99	.88	6	.98	.88
7	.99	.88	7	.85	.65
8	.99	.90	8	.93	.88
9	.89	.67	9	.98	.79
10	.96	.69	10	.98	.90
11	.99	.81	11	.93	.83
12	.80	.27	12	.91	.75
13	.84	.54	13	.70	.67
14	.96	.56	14	.95	.79
15	.93	.54	15	.72	.65
16	.96	.58	16	.84	.71
17	.80	.52	17	.79	.65
18	.60	.21	18	.91	.85
19	.82	.56	19	.88	.67
20	.71	.35	20	.86	.67
21	.89	.73	21	.81	.42
22	.91	.79	22	.79	.38
23	.91	.79	23	.46	.15
24	.60	.27	24	.72	.27
25	.78	.50	25	.77	.27
26	.82	.64	26	.79	.29
27	.29	.14	27	.63	.33
28	.51	.17	28	.84	.29
29	.67	.31	29	.60	.23
30	.58	.23	30	.60	.19
31	.80	.50	31	.72	.25
32	.27	.06	32	.56	.19
33	.76	.25	33	.70	.19
34	.53	.19	34	.74	.23
35	.49	.10	35	.74	.17
36	.49	.10	36	.40	.10
37	.13	.04	37	.58	.21
38	.27	.06	38	.72	.21
39	.33	.19	39	.53	.15
40	.44	.21	40*	--	--

---

*There was no item #40.

Appendix C
Item Difficulty List Continued

	Spelling			Arithmetic	
Item No.	Difficulty N-Inst.	Difficulty Inst.	Item No.	Difficulty N-Inst.	Difficulty Inst.
41	.18	.08	41	.84	.46
42	.42	.12	42	.86	.65
43	.31	.08	43	.88	.67
44	.44	.21	44	.88	.71
45	.31	.02	45	.72	.40
46	.13	.04	46	.81	.38
47	.02	0	47	.49	.44
48	.24	.04	48	.91	.62
49	.20	.06	49	.86	.54
50	.07	.04	50	.84	.52
51	.13	.02	51	.77	.35
52	.29	.08	52	.79	.29
53	.38	.14	53	.77	.42
54	.09	.04	54	.67	.29
55	.13	.04	55	.56	.23
56	.22	.04	56	.77	.40
57	.27	0	57	.56	.12
58	.33	.14	58	.40	.15
59	.51	.10	59	.51	.15
60	.07	.04	60	.70	.33
61	.24	.08	61	.53	.15
62	.29	.10	62	.60	.19
63	.13	.06	63	.35	.04
64	.22	.10	64	.33	.08
65	.33	.04	65	.56	.17
66	.07	.02	66	.58	.17
67	.24	.08	67	.53	.17
68	.18	.06	68	.33	.06
69	.04	.02	69	.60	.15
70	.13	.02	70	.30	.02
71	.22	.04	71	.53	.04
72	.36	.10	72	.42	.02
73	.47	.10	73	.07	0
74	.49	.10	74	.12	0
75	.33	.12	75	.23	.02
76	.18	.06	76	.14	0
77	.31	.10	77	.33	.04
78	.42	.12	78	.30	.04
79	.31	0	79	.42	.04
80	.02	0	80	.26	.02
Mean	.50	.28		.67	.37
Standard Deviation	.61	.30		.22	.28
r	.93			.87	

APPENDIX D

Mean Item Difficulties, Scores on Pilot Tests, CA, MA, and IQ for Each Main Study Subject (Grouped According to Treatment Condition)

Non-Institutionalized Subjects

Group	Subject Number	$\overline{DL}$	Arith. Score	Spell. Score	CA*	MA**	IQ**
A	020	.26	38	43	186	132	79
	019	.44	60	28	179	125	79
	031	.22	53	58	188	122	72
	022	.48	27	31	164	120	77
	010	.52	41	15	158	141	91
	021	.25	77	57	212	149	78
	044	.24	66	64	196	78	56
	011	.52	42	23	188	155	87
	005	.41	42	30	167	171	101
	014	.46	69	30	183	132	83
B	009	.17	25	32	172	151	93
	001	.14	40	67	171	195	112
	015	.18	74	39	188	132	83
	038	.18	65	32	221	127	70
	039	.17	54	26	182	101	68
	027	.14	27	53	175	113	75
	030	.12	49	56	197	122	72
	007	.16	53	25	172	167	96
	016	.18	77	27	187	141	81
	012	.15	77	62	197	151	85
C	003	.48	54	22	169	183	106
	034	.51	65	24	222	130	72
	040	.27	75	62	196	110	68
	017	.29	47	59	189	141	81
	013	.24	--	50	187	141	84
	002	.41	--	23	150	185	111
	041	.22	65	24	173	87	61
	032	.28	38	45	168	107	72
	043	.46	62	33	193	85	57
	045	.57	6	7	---	---	---

Appendix D Continued

Group	Subject Number	$\overline{DL}$	Arith. Score	Spell. Score	CA*	MA**	IQ**
D	028	.21	35	20	166	110	73
	024	.22	64	40	174	120	76
	029	.14	62	34	176	113	73
	006	.15	2	17	196	187	98
	004	.16	71	57	192	206	104
	008	.19	58	28	195	183	96
	035	.17	49	25	189	116	70
	026	.12	68	56	212	145	76
	042	.13	17	20	205	99	60
	033	.21	40	13	220	130	72
Mean		.27	48.35	36.42	186.03	135.97	80.72
Standard Deviation		.10	19.23	16.30	17.19	30.80	14.09

Institutionalized Subjects

Group	Subject Number	$\overline{DL}$	Arith. Score	Spell. Score	CA*	MA**	IQ**
E	128	.30	17	23	222	129	68
	139	.34	25	15	174	87	61
	125	.29	48	27	178	105	70
	113	.31	49	23	222	145	76
	120	.14	24	48	221	132	73
	116	.37	31	20	195	116	74
	121	.30	17	21	209	128	73
	102	.18	28	37	238	167	84
	110	.25	50	22	216	145	78
	115	.49	22	8	217	120	76
F	136	.08	37	21	206	103	62
	108	.05	62	66	181	125	79
	123	.08	35	34	177	105	71
	137	.03	66	63	223	105	61
	114	.08	66	32	220	144	76
	117	.13	71	26	211	130	74
	103	.06	12	47	193	149	84
	130	.12	11	20	177	105	68
	104	.11	20	19	183	132	83
	145	.08	30	15	230	97	57

Appendix D Continued

Group	Subject Number	$\overline{DL}$	Arith. Score	Spell. Score	CA*	MA**	IQ**
G	105	.30	19	20	234	160	82
	148	.34	6	16	167	69	53
	119	.33	29	23	226	132	73
	134	.45	15	12	221	120	65
	143	.22	27	27	202	94	57
	138	.48	3	10	218	105	61
	118	.39	30	18	185	113	73
	141	.50	13	9	185	112	60
	133	.51	12	7	231	148	65
	131	.26	43	28	181	97	66
H	107	.10	32	15	183	172	81
	124	.10	33	35	198	120	71
	109	.08	19	16	187	125	79
	146	.12	62	25	230	96	56
	127	.05	33	41	220	125	69
	112	.10	12	8	205	136	77
	111	.12	54	29	198	128	77
	106	.08	17	5	203	151	82
	132	.09	50	16	195	105	66
	126	.10	35	22	162	99	70
Mean		.21	31.62	24.22	203.10	121.90	70.78
Standard Deviation		.14	17.82	13.69	20.47	22.39	8.32

*at time of present study

**at time of last administration of Peabody Picture Vocabulary Test (see Appendix A)

APPENDIX E

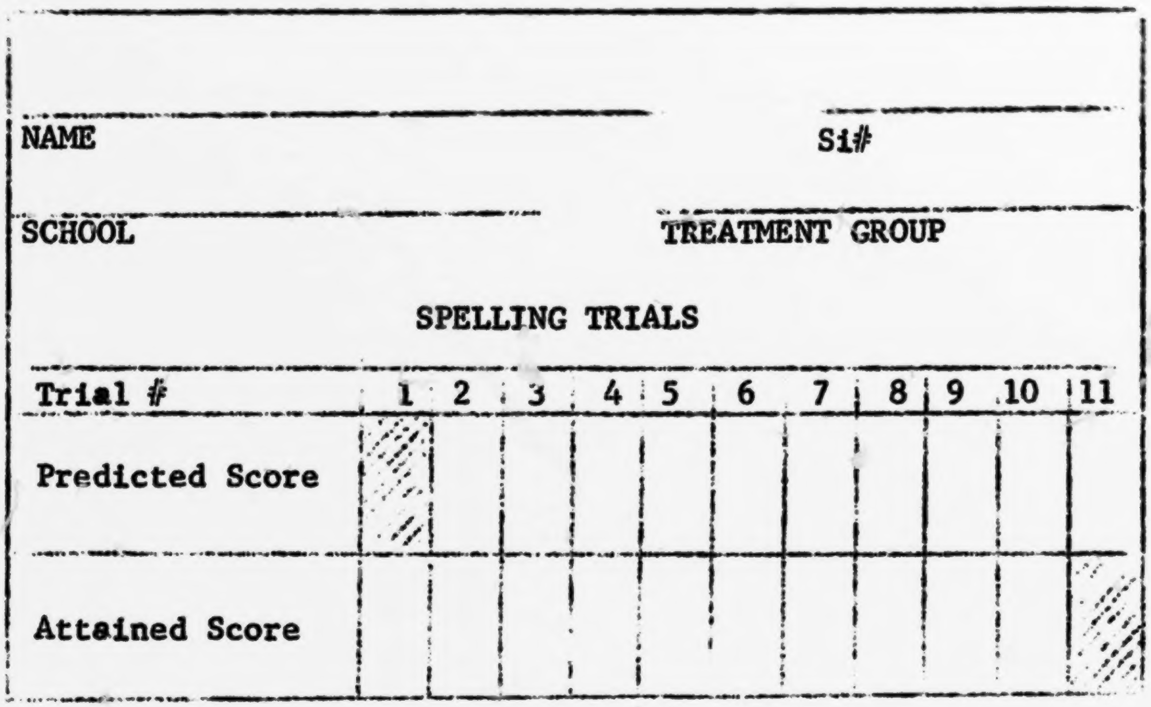
Sample of Stimulus Cards (Items) Presented to Main Study Subjects

DOCTOR

LESSONS

APPENDIX F

Sample of Record Form Used by Experimenters to Record Predicted Scores and Attained Scores for Individual Main Study Subjects



NAME ____________________ Si# ____________

SCHOOL ____________________ TREATMENT GROUP ____________

SPELLING TRIALS

Trial #	1	2	3	4	5	6	7	8	9	10	11
Predicted Score	▨										
Attained Score											▨

APPENDIX G

Raw Scores for Each Main Study Subject on Each Trial

Non-Institutionalized

Group	Trial Number									
A	1	2	3	4	5	6	7	8	9	10
020	6	7	6	8	8	9	8	9	9	9
019	1	3	4	4	5	6	4	5	4	4
031	3	6	7	10	10	10	(10)	(10)	(10)	(10)*
022	5	6	6	6	7	6	6	7	6	7
010	2	4	5	5	5	5	5	5	5	5
121	8	9	9	10	10	10	(10)	(10)	(10)	(10)
044	6	8	10	10	10	10	(10)	(10)	(10)	(10)
011	5	5	5	6	6	6	6	6	6	6
005	5	4	5	6	5	7	7	8	8	10
014	6	8	10	10	10	(10)	(10)	(10)	(10)	(10)
Sum	47	60	67	75	76	79*	76	80	78	81
Mean	4.7	6.0	6.7	7.5	7.6	7.9	7.6	8.0	7.8	8.1
B										
009	0	1	4	5	5	4	6	6	6	6
001	3	6	10	10	10	(10)	(10)	(10)	(10)	(10)
015	3	2	4	5	6	5	6	7	7	8
038	2	3	3	4	5	5	10	9	10	10
039	0	1	2	1	3	3	5	4	6	7
027	4	3	3	6	6	6	9	9	9	8
030	1	5	6	6	7	7	8	9	9	8
007	1	3	5	6	5	6	7	6	5	3
016	1	3	4	6	7	7	8	8	8	8
012	4	7	9	10	10	9	10	10	10	(10)
Sum	19	34	50	59	64	62	79	78	80	78
Mean	1.9	3.4	5.0	5.9	6.4	6.2	7.9	7.8	8.0	7.8

Appendix G Continued

Group	Trial Number									
C	1	2	3	4	5	6	7	8	9	10
003	5	5	5	6	6	6	6	7	6	6
034	3	4	4	6	7	6	6	6	7	8
040	6	9	9	9	10	10	10	(10)	(10)	(10)
017	6	9	9	9	10	10	10	(10)	(10)	(10)
013	5	4	4	7	10	9	6	10	9	9
002	3	4	3	6	5	7	6	7	7	7
041	5	6	7	6	8	8	8	7	9	9
032	4	6	6	6	6	6	6	7	7	7
043	5	6	8	8	8	9	8	10	10	10
045	1	1	3	3	2	3	3	3	3	3
Sum	43	54	58	66	72	74	69	77	78	79
Mean	4.3	5.4	5.8	6.6	7.2	7.4	6.9	7.7	7.8	7.9
D										
028	0	0	0	0	0	0	0	0	0	0
024	0	4	6	7	8	8	8	8	9	9
029	1	1	5	4	7	9	9	9	10	9
006	0	0	0	0	0	0	0	1	1	1
004	2	3	4	5	6	7	8	9	9	10
008	0	1	2	1	1	0	3	6	6	8
035	0	0	0	0	0	0	0	1	3	2
026	3	4	6	6	6	7	7	7	7	7
042	0	0	0	1	1	3	2	3	3	3
033	0	0	1	1	1	1	1	2	2	2
Sum	6	13	24	25	30	35	38	46	50	51
Mean	.6	1.3	2.4	2.5	3.0	3.5	3.8	4.6	5.0	5.1

Institutionalized

Group	Trial Number									
E	1	2	3	4	5	6	7	8	9	10
128	3	3	5	5	6	6	5	6	6	7
139	3	4	5	5	5	5	5	6	4	6
125	5	7	9	8	9	9	10	10	10	(10)
113	6	6	7	7	7	8	8	8	9	8
120	3	5	5	4	8	9	9	9	10	10
116	4	6	6	6	6	7	7	7	7	8
121	4	5	4	6	6	7	8	8	8	8
102	6	6	8	8	9	10	10	10	(10)	(10)
110	3	5	5	6	7	7	7	8	8	7
115	4	4	4	4	5	6	6	5	6	7
Sum	41	51	58	59	68	74	75	77	78	81
Mean	4.1	5.1	5.8	5.9	6.8	7.4	7.5	7.7	7.8	8.1

Appendix G Continued

Group	Trial Number									
F	1	2	3	4	5	6	7	8	9	10
136	0	1	2	3	5	3	2	5	4	5
108	3	8	10	10	10	(10)	(10)	(10)	(10)	(10)
123	1	5	5	6	8	8	8	8	8	9
137	3	9	9	9	10	10	10	(10)	(10)	(10)
114	0	2	5	5	6	7	6	6	7	7
117	1	2	2	4	8	9	9	9	10	10
103	1	2	5	4	5	5	4	7	6	8
130	1	1	4	3	2	3	3	5	4	7
104	0	1	1	1	2	3	1	2	1	2
145	0	0	0	0	1	0	0	1	1	0
Sum	10	31	43	45	57	58	53	63	61	68
Mean	1.0	3.1	4.3	4.5	5.7	5.8	5.3	6.3	6.1	6.8
G										
105	5	7	7	5	6	7	8	6	7	7
148	4	4	4	4	4	4	4	4	4	4
119	4	6	7	8	7	9	8	8	9	10
134	4	4	5	5	5	5	6	5	5	5
143	4	6	8	8	7	8	8	8	8	10
138	5	5	6	5	5	5	5	5	5	5
118	4	5	6	6	5	6	6	6	6	7
141	5	5	5	5	5	5	5	5	5	5
133	5	5	5	5	5	5	6	6	6	6
131	7	8	10	10	10	(10)	(10)	(10)	(10)	(10)
Sum	47	55	63	61	59	64	66	63	65	69
Mean	4.7	5.5	6.3	6.1	5.9	6.4	6.6	6.3	6.5	6.9
H										
107	0	0	0	1	1	1	1	1	1	2
124	0	3	5	6	6	5	5	7	9	9
109	0	2	2	1	2	3	1	2	3	4
146	2	5	7	7	7	8	10	10	9	10
127	4	4	5	6	7	9	8	8	9	10
112	1	1	1	1	1	1	1	2	2	3
111	2	2	2	4	3	3	3	4	4	5
106	0	0	0	0	0	0	0	0	0	0
132	0	0	0	1	1	1	2	2	2	3
126	1	1	2	2	4	4	4	3	4	4
Sum	10	18	24	29	32	35	35	39	43	50
Mean	1.0	1.8	2.4	2.9	3.2	3.5	3.5	3.9	4.3	5.0

---

*Scores in parentheses indicate that S attained 3 successive errorless trials before reaching trial #10. For computational purposes it was assumed that S would have continued to score 10 had he continued tasks after 3 successive perfect trials.